Precautionary Motives and Portfolio Decisions

by

Stefan Hochguertel*

Tilburg University†

June 1997

Abstract

Theory predicts that under certain restrictions on preferences prudent consumers will allocate relatively more funds to riskless assets when there is uninsurable background risk. This paper analyzes empirically the relevance of precautionary motives for the structure of household wealth. To this end, a new and rich data source from the Netherlands is exploited. The question of primary concern is: *what impact, if any, does the presence of income uncertainty have on the structure of Dutch households' portfolios?* We employ various semi–parametric estimators, both for cross–sections and for panel data to assess the response of households' portfolios to uninsurable background risk. We find some, but not unanimous support for the view that portfolios become less risky as income uncertainty increases.

Keywords: precautionary saving, background risk, household saving, portfolio choice, application of LDV models

JEL classification: *D12, E21, G11*

*The author is supported by the VSB savings project. Thanks to the VSB–CentER savings project for providing the data. I would like to express my gratitude to Rob Alessie, and Arthur van Soest for discussions & helpful suggestions. Thanks for comments are due as well to Erwin Charlier and seminar and conference participants at Tilburg and the TMR workshop on "Savings, Pensions, and Portfolio Choice" at Naples, February 1997, and Agar Brugiavini in particular. Thanks to Bernd Fitzenberger who provided the FORTRAN code of his algorithm. Errors are and will remain mine.

†correspondence address of the author:
CentER for Economic Research, Tilburg University, PO Box 90153, 5000 LE Tilburg, The Netherlands; phone: +31 13 466.3334; fax: —.3066; e–mail: stefan@kub.nl.

1 Introduction

This paper studies the impact of income uncertainty and precautionary motives on the structure of households' portfolios. The theory of precautionary savings (Leland (1968) and Kimball (1990)) suggests that, given certain properties of the underlying preference structure, individuals facing uninsurable background risk will accumulate additional resources to buffer adverse income shocks. Not only the level of wealth accumulation will be affected, but the additionally generated savings will be channeled into certain assets: households will allocate a larger share to riskless or liquid assets (Kimball (1992, 1993)) while the demand for risky assets will decrease as background risk increases (Koo (1991) and Bertaut and Haliassos (1992)).

This topic is important since it deals with responses to uninsurable risks, ie. risks which due to incompleteness of insurance markets cannot be shared among agents. One might argue (as Allen and Gale (1994) in their introduction) that even though financial innovations help to remove incompletenesses in financial markets, the importance of uninsurable earnings risk is increasing because employment relations become less stable over time. Then, handling of earnings risk becomes more important. To better understand how portfolio risks in the presence of unrelated risks are allocated in financial markets, a first step is to investigate the portfolio response of individual households to this type of risk, which we assume to be uninsurable. The current paper investigates whether empirical evidence can be found to support the theoretical predictions.

We study Dutch panel data. The Netherlands are an interesting case for studying precautionary saving because the country's national saving rate (14%, 1988) is quite high in international comparison, while at the same time mortgage constraints like down payment requirements are absent (Hochguertel and Van Soest (1996)). Thus, prima facie, liquidity constraints would presumably not play an as important role to explain saving as they do in other countries. Another piece of prima facie evidence is that around three quarters of interviewed respondents in our data would assign a value of 5–7 on a 7–point scale to the statement: Is it to you personally of much [7] or little [1] importance: Saving *"as a reserve to cover unforeseen expenses"*. Since most non–housing wealth of Dutch households is held in saving related assets, and the household stock–ownership rate in the Netherlands is one of the lowest among advanced economies (Hochguertel et al. (1997)), this raises the question whether earnings uncertainty induces households to hold more riskless assets, provided there is a precautionary motive.

The empirical analysis in this paper is based on estimating the demand for riskless assets in the form of a budget share equation. We use a new survey on Dutch households (VSB panel), which was specifically targeted at the wealth structure and savings behavior of households and thus provides an excellent basis to address wealth and savings related questions. The survey constitutes a rich source of household wealth information and also comprises measures of individual attitudes to saving and income expectations. Several possible choices for an operational measure of income uncertainty and precautionary motives from the expectations data are discussed.

Some emphasis is put on simple models which are estimated with robust methods. This point is important since theory does not yield closed form solutions for the equations to be estimated. For cross–section data we use both parametric models and Powell's (1986a) censored regression quantile estimator; the latter is distribution free and robust to unknown forms of heteroskedasticity. Especially the differences across quantiles prove to be of interest. Additionally, we exploit the panel nature of the data, applying Honoré's (1992) semiparametric Tobit estimator for fixed individual effects. These fixed effects represent time–invariant unobserved preference heterogeneity, for example. Incorporation of fixed effects has been missing so far in the empirical analysis of the impact of precautionary motives on household portfolio choice. We apply the selected semiparametric estimator to both balanced and unbalanced panel data.

We explore the robustness of the findings by extensive scrutiny with respect to model specification. The analysis reveals that there is some, but not unanimous, support for the view that income uncertainty increases the share allocated to riskless assets.

The remainder of the paper is organized as follows: in the subsequent section we specify more precisely the implications for portfolio decisions in the presence of background risk, and provide a brief overview of some related papers. The third section gives a description of the data and discusses measures of income uncertainty and precautionary motives. Section four outlines the econometric models we consider, section five discusses the findings. The final section concludes.

2 Theoretical Implications & Empirical Findings

2.1 Precautionary Motives and Portfolio Choice

Leland (1968) was the first to point out a motive for precautionary saving when the utility function exhibits convex marginal utility. Kimball (1990) introduced the notion of prudence in analogy to the Arrow–Pratt measure of (absolute) risk aversion and showed that the negative of marginal utility plays the same role for precautionary saving as the utility function for risk aversion: concavity of the utility function $U(U'' < 0)$ implies risk aversion, concavity of $-U'(U''' > 0)$ implies prudence. Kimball proposes to measure the degree of prudence as $\pi = -U'''/U''$ which indicates the strength of the precautionary saving motive. In the presence of background risk, decreasing absolute risk aversion (DARA) together with positive prudence imply positive precautionary savings. Kimball also introduces the concept of a precautionary premium which is the amount of money that must be given to a risk averse consumer to compensate her for unchanged behavior as the uncertainty of the environment is changed. He shows, that under DARA the precautionary premium exceeds the risk premium, meaning that the precautionary motive is stronger than risk aversion. This had previously been established by Drèze and Modigliani (1972). Kimball (1992) shows in a two–period model that a risk which is compensated such that expected utility stays the same, will still raise saving since it increases the second–period marginal utility. An implication is that higher savings make it more desirable to take on compensated risks. On the other hand, this complementarity between savings and compensated risks implies that a compensated risk makes saving more attractive as well. All that is said so far just means that prudent consumers cope with risks (which are otherwise uninsurable) by accumulating more life–time resources than they otherwise would.

But the question remains: what determines how the additional funds are allocated across assets? Kimball (1993) clarifies the implication of prudence for asset demand. He shows that "standard risk aversion" (defined as global decreasing absolute prudence (DAP), or local DAP with DARA) is sufficient to decrease the level of investment in a risky asset in the presence of an independent zero–mean background wealth. Kimball (1992) introduces the term "temperance" (often measured as $\tau = -U''''/U'''$) which describes a desire to reduce total exposure to risk. This involves that in the presence of one risk, the decision maker tries to avoid other risks. Thus, there is a negative interaction between those risks, if they are statistically independent. The implication for a portfolio choice problem is that temperance induces individual investors to allocate relatively more funds to riskless assets. Under standardness, temperance is greater than prudence, which in turn is greater than risk aversion, which is positive ($\tau \geq \pi \geq \alpha > 0$). Prudence leads to more precautionary savings, while temperance induces the investor to allocate relatively more to riskless assets, it can even lead to absolutely lower holdings of risky assets. Koo (1991) confirms this for a multiperiod portfolio model.

The upshot of the discussion is that the portfolio effect of background risk depends on whether certain restrictions on the utility function hold. This is mainly an empirical question.

2.2 Wealth Elasticities and Risk Attitudes

The simplest way to look at portfolio choice problems is to analyze a two asset economy, one asset being riskless, the other one risky. The traditional approach in expected utility theory has established the relationship between wealth elasticities and risk aversion. Under specific circumstances, one can even say something about precautionary motives and portfolio choices, as we will show now:[1] consider the problem of maximizing expected

[1] The roots can be traced back to the seminal contributions of Arrow and Pratt; the exposition here follows Eeckhoudt and Gollier (1995).

utility of (random) final wealth w_1, $\mathrm{E}(U(w_1))$. Final wealth is composed of investment in and returns to a risky asset a, and a safe asset $m = w_0 - a$. w_0 is initial wealth. Thus, $w_1 = w_0(1 + r_m) + a(r_a - r_m)$. Note that the first order condition

$$\mathrm{dE}(U(w_1))/\mathrm{d}a = \mathrm{E}(U'(w_1)(r_a - r_m)) = 0 = g(a; w_0, r_m, F_{r_a}(\cdot)) \tag{1}$$

is a function $g(\cdot)$ of the choice variable a, and parameters such as the riskfree rate, $r_m > -1$, initial wealth, w_0, and the distribution of the risky rate of return, $F_{r_a}(\cdot)$. Expected utility is uniquely maximized (SOC) under the assumption of global risk aversion, $U'' < 0$:

$$\mathrm{d}^2\mathrm{E}(U)/\mathrm{d}^2 a = \mathrm{E}(U''(w_1)(r_a - r_m)^2) < 0. \tag{2}$$

To analyze the effect of w_0 on a, differentiate g totally to obtain

$$\mathrm{d}a/\mathrm{d}w_0 = (\partial g/\partial w_0)/(-\partial g/\partial a). \tag{3}$$

Since the denominator of (3) is positive, the sign is that of the numerator,

$$\mathrm{E}(U''(w_1)(r_a - r_m)). \tag{4}$$

In order to determine the sign of (4), one needs to determine the degree of absolute risk aversion, $\alpha(w_1) \equiv -U''(w_1)/U'(w_1)$:

$$\mathrm{E}(U''(w_1)(r_a - r_m)) = -\mathrm{E}(\alpha(w_1)U'(w_1)(r_a - r_m)) \tag{5}$$

is positive under DARA,[2] and risky assets are a normal good. In order to see whether a is a luxury under DARA, further assumptions about the degree of relative risk aversion, $\beta(w_1) = w_1 \cdot \alpha(w_1)$ are needed. From (3) write the elasticity of the demand for risky assets with respect to initial wealth as

$$\eta_{w_0} = -(\partial g/\partial w_0)/(\partial g/\partial a) \cdot w_0/a$$

Using (4) and (2) we can re–write this as

$$\eta_{w_0} = \frac{-(1 + r_m)\mathrm{E}(U''(w_1)(r_a - r_m))}{\mathrm{E}(U''(w_1)(r_a - r_m)^2)} \cdot \frac{w_0}{a},$$

or

$$\eta_{w_0} = 1 + \frac{-(1 + r_m)w_0\mathrm{E}(U''(w_1)(r_a - r_m)) - a\mathrm{E}(U''(w_1)(r_a - r_m)^2)}{a\mathrm{E}(U''(w_1)(r_a - r_m)^2)} \tag{6}$$

Given positive demand for risky assets ($a > 0$), the numerator in (6) is decisive. It can be written as

$$-\mathrm{E}(U''(w_1)(r_a - r_m)w_1) = \mathrm{E}(\beta(w_1)U'(w_1)(r_a - r_m)).$$

[2] To understand this, note that w_1 and r_a are positively correlated, given r_m, w_0 and a. Hence, $r_a > r_m$ when w_1 is "large". Under DARA, $\alpha'(\cdot) < 0$ and thus $\alpha(w_1)$ weighs $U'(w_1)(r_a - r_m)$ more heavily if $r_a < r_m$ than if $r_a > r_m$. The expected value in the right hand side of (5) will then be negative.

Under the common assumption of CRRA, $\beta =$ const. can be taken outside of the expectation and hence $\eta_{w_0} = 1$ due to the FOC (1). Under DRRA, it follows that $\eta_{w_0} > 1$.[3] In words: the investor is willing to give up a smaller fraction of her wealth in return for insurance against a multiplicative risk as wealth increases. Therefore, a larger fraction is invested in the risky asset with wealth. The relationship between the wealth elasticity of the demand for risky assets and risk attitudes is summarized in table 1 (Appendix B).

Recall from the definition of absolute risk aversion (α) that absolute prudence can be written as $-U'''/U'' = \alpha - \alpha'/\alpha$. From here, it follows directly — after some simple algebraic manipulations — that CRRA implies DARA and DAP (standardness, Kimball (1993)). From a consistent estimate of the wealth elasticities of risky assets one could possibly infer whether background risk will have an impact on the portfolio allocation, at least if $\eta_{w_0} = 1$ holds. But if $\eta_{w_0} > 1$, for instance, this is not clear anymore: in this case, we have DRRA (implying DARA) but cannot exclude the possibility that DAP does not hold.[4] Therefore, an empirical analysis calls for models which relate portfolio allocation to background risk.

2.3 Related Empirical Literature

The empirical literature on precautionary savings has one of its roots in Euler equations models testing the Life Cycle Hypothesis (LCH).[5] What those studies discovered was that the data could not sufficiently be explained by theoretical predictions: consumption has been found to be excessively sensitive to transitory income innovations to be in accordance with the LCH. Moreover, it was found that the elderly have a tendency not to dissave, which could be due to precautionary saving (accidental bequests, Davies (1981)). Most of this literature was based on macro data, ignoring idiosyncratic income variability. Zeldes (1989) showed that one would need to go beyond the case of certainty equivalence utility functions in order to obtain consumption paths which could reconcile some of the empirical regularities. Zeldes' paper made a strong case for precautionary saving, using household data.[6]

With the use of micro data it became customary to derive some measure of income uncertainty from an estimated income process (see for instance, Carroll and Samwick (1995a,b), Hubbard et al. (1994), or Jianakoplos et al. (1996)). Kazarosian (1997), for instance, estimates a specification which explains the current asset–to–permanent income ratio from income uncertainty and permanent income. Both permanent income

[3] The conclusion follows by similar reasoning as for (5), cf. fn. 2.

[4] The condition for DAP to hold is that the second derivative of RA, α'' is either positive, or, if it is negative exceeds $\alpha\alpha'[1 + \alpha'/\alpha^2]$, which is negative since DRRA implies DARA implies that prudence exceeds RA (Kimball (1990)). This condition is stronger than risk vulnerability (see Gollier and Pratt (1996)) or even properness (Pratt and Zeckhauser (1987)).

[5] For introduction and overview see Deaton (1992).

[6] More can be found in the excellent survey of Browning and Lusardi (1996).

and income uncertainty are derived from estimates of a random effects (RE) model for current income, explained from occupation and age. Focus is on explaining observed behavior in the 1966 wave, incorporating information on permanent income and income uncertainty which thus pertains to future values. Two uncertainty measures are considered and yield significantly positive estimates.

Based on this approach, Chakraborty and Kazarosian (1996) study portfolio composition by regressing the share of specific assets in total wealth on occupation and employment variables, among others. Distinguished asset types are risky and illiquid assets (housing, real estate, business), safe and liquid assets (government bonds, savings accounts) and risky and liquid assets (shares, bonds, mutual funds). The authors find that people in general move away from assets which are both risky and illiquid, and towards liquid assets, both risky and safe, as income uncertainty increases.

The idea of estimating income processes is a viable approach to assess income uncertainty in the absence of other information relating to future income developments. It is important to know what people expect their income to be in the future and how certain or uncertain they are about their own expectation. A variation in income over time only bears a loose connection to this requirement. In general, consumers themselves will have more information to predict their own income than an outside observer who has to rely on historical income data. In this sense, it is preferable to use subjective expectations data about income and income variations.[7] The Italian SHIW data contain some information about this, pertaining to a 1 year horizon. The SHIW data were used by Guiso et al. (1992, 1996) to assess the relevance of income uncertainty for precautionary saving, and portfolio choice, respectively. To correct for biases stemming from the general price development, the authors can also exploit information about inflation uncertainty. The reported findings attribute only a small role for income variance in saving behavior. Their measure might only capture transitory income shocks, and be of more concern for liquidity constrained households. Lusardi (1996) re–assesses the Guiso et al. measure and interprets it as dislike of income uncertainty on the background of job stability. She reports evidence that the probability of a non–zero income variance falls with on–the–job tenure. She finds a somewhat stronger impact for income uncertainty on saving than Guiso et al. (1992).

Guiso et al. (1996) analyze the impact of human capital risk on financial portfolios. They model the choice between risky and non–risky assets by a two–limit Tobit specification. The definition of risky assets encompasses equities, investment fund units, and corporate and government debt. An alternative, broader measure includes saving accounts as well. As saving accounts dominate financial assets, their inclusion in the category of risky assets does have a major impact on the distribution of the dependent variable, but

[7] But note the finding of Das and Van Soest (1996) that heads of households tend to systematically underestimate their income growth; this calls into question rational expectation formation.

the coefficient on income variance in the regression is hardly affected. Expected inflation variance is insignificant in all specifications but has a noticeable impact on the magnitude of the coefficient of income variance. The authors also investigate the impact of liquidity constraints on portfolio choice and find a significantly negative value only if saving accounts are included in the definition of the dependent variable. In a further robustness analysis, the size of the coefficient on income variance turns out to be sensitive to the assumptions of homoskedastic normal errors of the Maximum Likelihood Tobit estimator.

The current paper compares closely to that of Guiso et al. (1996) in many respects, since both use micro–data on assets, model portfolio choice in comparable ways, and assess income uncertainty from subjective information. In our paper we put somewhat more stress on robust estimation, but deviate from Guiso et al. in a couple of more essential respects: we discuss in more depth alternative measures of income uncertainty and compare the robustness of the findings across those different measures, and, next to cross–section data we also employ panel data information. The latter has the advantage that we can accommodate unobserved individual effects which can capture important behavioral characteristics.

Both the interesting data source and the econometric techniques employed merit some further discussion. We will next describe the data we have available. As it turns out, we have information which allows us both to avoid arbitrary assumptions on the underlying income process and to infer impacts from long–term uncertainty.

3 Data, Variables, and Summary Statistics

3.1 Data Set

For the question addressed in this paper we favor a micro data set which entails detailed information on the asset structure of household wealth, and which allows to assess idiosyncratic income risk or precautionary motives of individual decision makers.

The data set we choose is the so–called VSB panel, launched by the VSB–CentER savings project in 1993. For cross–section analysis, we exploit the first wave, containing flow (income and work) information for 1992 and wealth related information for January 1, 1993 (1993 wave). Analogously, the 1994 wave provides flow information for 1993 and stock information for January 1, 1994, and so on. The information for this second wave has actually been requested only a few months after the information for the first wave had been collected. Therefore, many of the questions were only asked to the new entrants to the panel and not to old households. A second complete wave is again available for 1995.

The VSB panel consists of two samples, one representative of the Dutch population (REP),[8] covering approximately 2000 households, and one sample drawn from the upper

[8]As it turns out, home owners are over–represented in the "representative" sample; this might have

decile of the income distribution (HIP),[9] encompassing roughly 800 households. The latter has the advantage of yielding more insight into financial behavior of richer households due to the higher asset level, and more diversified portfolios. The data collection has been performed on–line via terminal sessions. Respondents were asked to key in answers to questions they received via a modem on a personal computer. The data set is in its features most closely comparable to the triennial US Survey of Consumer Finances.

Compared to other data sets, the VSB panel has considerable strengths. The questionnaire has been specifically tailored to facilitate micro–econom(etr)ic and psychological research on savings, in a very encompassing way. Questions covered general information on the household and individual household members, and detailed questions on income from all sources, work history and current work status, pension claims, health, and assets and debts. Respondents were requested to look up income and wealth information from pay slips, tax forms, and account statements, as far as possible. Ownership information on assets and debts is complete and known even for those cases where the amount held was unknown. In addition, there is a set of questions referring to economic–psychological concepts which in principle are designed to allow inferences about time preference rates, risk aversion, financial attitudes and savings behavior, and choice of banks. The current paper employs variables from all sections of the questionnaire, although particular emphasis is put on assets and debts and the subjective information from the "psychological" part.

3.2 Variables

"The main difficulty in the empirical analysis is to find appropriate measures of income risk [. . .]" (Guiso et al (1996, p. 160)).

To analyze the impact of income uncertainty on portfolio composition we explain the (budget) share allocated to certain assets in total assets by a number of household characteristics, wealth, and some measure of income uncertainty or precautionary motives.

The wealth information is available at the asset level per respondent. We aggregate over assets per respondent and over respondents within each household. RHS variables relate to the information for the head of the household.

As dependent variable we choose the ratio of riskless and liquid assets to financial wealth. We select checking accounts, saving accounts, and deposit books into the category "riskless and liquid". We include neither liquid, but risky assets, nor riskless but illiquid assets. Other available assets (employer sponsored savings plan, savings certificates, growth funds, mutual funds, bonds/mortgage bonds, shares, life insurance contracts) de-

an effect on the portfolio structure and wealth holdings of households in general.

[9] the cut–off point has been chosen to be 105,000 guilders

fault into residual financial assets.

In terms of precautionary savings any asset can be used to cover expenses due to unforeseen circumstances, such as particularly bad draws in earnings, even if there are considerably high transaction costs associated with it (see Carroll and Samwick (1995b) for discussion.) The choice made here includes those assets which can broadly be viewed as “riskless”. Housing, other real estate, and mortgages, are often viewed as risky assets as well (cf. Henderson and Ioannides (1983)), and would enter the denominator of the dependent variable when we would use total wealth. Here, we observe a substantial number of households with negative total net worth due to high mortgages, violating the interpretation of the dependent variable as a budget share. Moreover, in the second wave of the survey (1994), real estate wealth is only available for new panel entrants.

For the same reason why we do not consider total net worth, we exclude negative checking account balances and consumer debts. Thus, the budget share as the dependent variable is censored at zero and one. Observations with zero financial wealth drop out. This concerns 12% of the observations. In addition, 23 observations drop out due to missing values on income uncertainty. In the remaining sample, 2.2% of the observations are censored at 0 and 43% at 1.[10]

Both riskless and financial wealth are the sums of several asset components, each of which is affected by item–nonresponse. A missing value in one of the components carries through to the aggregate and thus causes severe problems of missing observations. In order to reduce this, the missing values have been treated as zeros, although they are known to be non–zero. Thus, the obtained wealth measures are for some observations “minimum levels” — the true asset holdings for those households are larger and thus the wealth measures are censored from below at that minimum level. Clearly, when the dependent variable is a share where both numerator and denominator are subject to this censoring problem, this causes an intricate problem which is not straightforward to solve. For an equation explaining the level of wealth this would lead to a specification where censoring is from above at individual censoring points (i.e. at the “minimum levels”). But for the share equation we will thus ignore the separate censoring on nominator and denominator of the share.[11]

In the model which we consider, we will explain the portfolio share of riskless and liquid assets from household characteristics like age, education, financial wealth, labor market variables, marital status, some income indicator, and some measure of income uncertainty or precautionary motives. Variables such as age and marital status are self–explanatory. Explanations for some other RHS variables are given below in the note to

[10]These numbers refer to 1993 and the baseline specification we will focus on.

[11]One can straightforwardly take into account censoring on the numerator only. This then translates into censoring on the share as additional right–censoring (3.5%) with individual censoring points. This kind of “censored regression model” is rather ad–hoc and estimates do not differ substantially from Tobit estimates.

table 4. We focus on measures of income uncertainty.

Before introducing the alternative measures, we discuss an important point. The theoretical literature on background risk and portfolio choice does not address the question which evaluation horizon is important for assessment of income uncertainty. This is because models are usually of an atemporal type. It has been argued, however, that it is life–time uncertainty which matters for the level of precautionary saving (see, among others, Carroll and Samwick (1995a,b)). So, if consumers optimize their consumption paths by maximizing expected lifetime utility, the whole future process of income risk will be relevant for that consumption plan, and hence, saving.[12] Applied to the portfolio choice problem, this argument would favor a long–run measure of uncertainty. Admittedly, this is not so clear, as portfolios can be re–structured even in response to short term shocks to income, without having major repercussions for the level of saving or the expectation of the consumption path in the long–run. The reason why this is not frequently observed might be entirely due to transaction costs. In the analysis we will both estimate cross–section and fixed effects panel data models. In panel data models, transaction costs which are time–invariant can be captured by the fixed effects, such that the distinction between long and short horizon becomes less important.

If we assume income uncertainty to be generated from exogenous earnings processes, we can use some measure of it as a regressor. Care has to be taken in constructing such a measure. Two possibilities exist:

1. An ordinal measure which indicates to which extent heads of households are certain about their own expectations regarding household net income; this measure is available for a short (12 months) and a longer (5 years) horizon; for reference, this measure will be called **"income uncertainty"** in the sequel. In cross–section models, the longer horizon measure will be interpreted as the analogue of "permanent income shocks". The short–run measure should have an impact if liquidity constraints are of importance. We will weigh up the short–run against the long–run measure in a sensitivity analysis. The information in the data set is a categorical variable for four different degrees of uncertainty about the household's income expectation, and conditional on a point expectation for income growth. This variable is split up into two dummies indicating whether the head of household is either "rather certain" (*moderate* income uncertainty), or "not very certain" or "not certain at all" (*high* income uncertainty). The pooling of two categories seemed called for due to the relatively low frequency (2.6%) of the last category. The reference category is "very certain" (*low* income uncertainty).[13]

[12]See Blundell and Stoker (1994) for an analysis of the impact of the timing of income risk on consumption and consumption growth. It appears that assumptions about the particular form of risk aversion are important for the timing effect.

[13]For wording of the survey questions cf. Appendix A, variables (c) and (g)

2. A continuous measure of the variance of expected income one year ahead, constructed from a set of questions on the likelihood of given income changes. The answers to this set of seven questions were ordinal but have been converted into a continuous measure of variance under some simplifying assumptions.[14] The measure is roughly similar to the one chosen by Guiso et al. (1992, 1996) for Italian survey data. It will be called **"income variance"** from now on.

Alternatively, we can consider direct measures of precautionary motives:

3. A range of questions asks directly for saving motives. Heads of households were asked to indicate on a scale from 1 to 7 the degree of importance for a number of saving motives. Among them are motives for saving "for unforeseen expenses as the result of injuries or accidents", "to cover income losses due to unemployment", "for unforeseen expenses", and "to cover outstanding debts". We refer to them simply as **"precautionary motives"**. The original precautionary motive is probably best captured by the general motive for unforeseen circumstances, but as all these motives are related, it might be sensible to summarize them in one measure, applying factor analysis. The information for these variables is available for the 1993 and 1995 waves and only for new panel entrants of the 1994 wave.

For our cross–sectional estimates, we regard a long–run horizon as more appropriate to capture permanent income uncertainty. This information is only available for the first measure, and at best for a five year horizon. We assume that a five year horizon is a good approximation to long–run uncertainty. At any rate this horizon is far longer than the evaluation horizon in other data sets. If it were extended much further into the future, the issue of possible endogeneity of income risk would have to be discussed. We also assume that two different individuals with the same uncertainty would classify themselves in the same category. Thus "rather certain" means the same degree of uncertainty across individuals. As the "income uncertainty" measure is ordinal, calculation of elasticities is precluded.

The second measure is continuous, but we have to make some simplifying assumptions in order to retrieve it from the data. It pertains to income uncertainty 1 year ahead. Note, that a measure of variance might not be suitable for investigating the impact of background risk on saving and portfolio behavior since in general variance is not a sufficient statistic of risk. An exception is a CARA utility with additive and normally distributed income shocks (cf. Carroll and Samwick (1995b)).[15]

[14]The variables involved and the a more detailed description of the construction of this measure, including the assumptions made is deferred to Appendix A.

[15]Instead, measures like the integral under the cdf of the random variable causing the risk or a measure based on transfers of probability weights have been proposed, but they cannot be constructed from our data (Rothschild and Stiglitz (1970)). Carroll and Samwick (1995b) advocate a measure which mirrors Kimball's (1990) equivalent precautionary premium.

Note that the corresponding questions in the survey refer to "income" which might include asset income as well. Asset income uncertainty can more easily be avoided than uncertainty from labor income, and thus the obtained measure of uncertainty might be contaminated by portfolio risk. Given the low share of high–return assets in household portfolios, the asset income will not be substantial for most households, though. Moreover, even a measure of 'pure' labor income risk might not be exogenous as people have a choice of occupation and labor supply (see for instance, Lusardi (1996)).

Both these measures are based on "household net income". This income concept has a couple of important advantages: first, as it is after–tax, it does not have to be corrected for the insurance effect of taxation, which as such can lead to higher demand for risky assets (Elmendorf and Kimball (1991)). Second, as it pertains to the entire household, possible risk–pooling within two or multiple earner households is already taken into account. We are left with uncertainty that will not be influenced anymore by tax effects or income pooling.

The setup of the questionnaire does not explicitly state that real income changes should be considered; thus, the obtained measure for income uncertainty might have to be corrected for the impact of inflation. Although there are questions which relate to the inflation expectations,[16] both one and five years ahead, there is no information which would allow to construct a measure of the second moment of the distribution of expected inflation; thus, the approach as outlined in Guiso et al. (1992, 1996) is not feasible. Instead, in the sensitivity analysis to follow we use a direct measure for the (level of) price expectations of 5 years ahead.[17]

The "saving motives" as discussed above, serve to control directly for precautionary motives. The questionnaire does not specify an evaluation horizon. In a way, these variables measure properties of the preference structure. Income uncertainty will only have an impact on saving behavior if households actually perceive it as a source of risk that they have to forearm against by taking appropriate action (resource accumulation and risk shifting), and if it is properly measured. A second interpretation is to view them as proxying consumption risk, since the survey questions refer to "expenditures".

All the discussed measures have the advantage that they are readily available and no imputation has to be done. In particular, they are forward looking and need not be based on historical information. From the discussion above, we tend to prefer the "income uncertainty" measure. The other two are discussed briefly when we consider sensitivity

[16] variables $(h) - (k)$ in Appendix A

[17] Guiso et al. (1996) report insignificant parameter estimates for inflation variance. Since one might argue that in the Italian case inflation is both on a higher level and more volatile than in the Dutch case, we have a prior that biases due to omitted inflation variance are negligible. Further support derives from Das and Van Soest (1996) who make a case that Dutch households are concerned with real income when asked questions about future income expectations where the questionnaire does not explicitly specify this. Their measure is based only on a 12 months horizon.

of the results with respect to the specification in section 5.

3.3 Summary Statistics

This subsection presents selected summary statistics of the data. Table 2 contains cross–tabulations for the "income uncertainty" measure across evaluation horizons and across panel waves, table 3 shows a breakdown of the share of riskless assets and of income uncertainty by exogenous variables for 1993. We show the figures for those observations which are used subsequently in the econometric analyses.

Tables 2a/b cross–tabulates the answers to the degree of income uncertainty[18] conditional on the expected income change,[19] for the year 1993. There seems to be higher uncertainty for people expecting their income to change over time. People expecting the same income next year are relatively more certain about this. In the short run, labor contracts are fixed and income can be predicted quite well. This pattern can be observed for the longer run as well, though. There is a higher percentage of people who are certain among those who expect their income to fall over next 5 years, and a lower percentage among those who expect an increase. There are about twice as many people who expect their income to rise over the next five years than people who expect a drop. Income uncertainty rises[20] as the horizon is shifted outward. This is documented in table 2c. Especially the first category loses, whereas the third gains. Otherwise, both measures are strongly correlated. In addition, there are some fluctuations across sampling years, as can be seen from tables 2d, but they are rather unsystematic. We would therefore not expect a large year effect.

Turn now to table 3. The first two columns display some of the most important right hand side variables in classes and the percentage of the respective class in the sample. Columns three through seven describe the distribution of the share of riskless assets (means and selected percentiles). The last three columns of table 3 contain the three different categories for income uncertainty, with a 5 year horizon. Here, the percentages add up rowwise to 1.

We observe a u–shaped age pattern for the mean share of riskless, liquid wealth: whereas the lowest share is observed in the age group of 40 to 54 years, both younger and older households invest a larger share in riskless assets. This pattern also obtains for the 10th–40th percentile. In the youngest age group, more than 70% of households do not hold any but liquid, riskless assets. Stratifying by education reveals that university graduates invest relatively least of their financial wealth holdings in riskless assets. The lowest education group holds a more riskless portfolio.[21] Self–employed hold somewhat

[18]the exact wording is given in Appendix A, variables (c) and (g).
[19]variables $(a-b)$ and $(e-f)$, Appendix A.
[20]in terms of higher frequencies for the category "high"
[21]cf. notes to table 4 for a definition of education levels.

less liquid and riskless portfolios; but still, between 60% and 70% of all self–employed hold more riskless than risky assets. Particularly liquid and riskless portfolios are held by the unemployed. They play a minor role in the sample, however. A monotonic pattern arises for financial wealth: the higher financial wealth, the lower the share of riskless assets. There is a substantial difference between members of the representative panel and high income households. As expected, richer households hold a lower share of their financial wealth in riskless and liquid assets. Splitting by income uncertainty reveals those households who are certain about their income expectation to hold relatively more risky portfolios than others. The pattern for the share of riskless assets is monotonically increasing in income uncertainty, as theory predicts.

We will now discuss income uncertainty across subgroups of the sample. Splitting by age reveals that people in the typical pre–retirement age (55–64 years) seem to face rather low income uncertainty; they would typically know their position in the labor market and have a rather distinct picture of their income development. For the youngest age group income uncertainty is rather high. This might reflect career uncertainty. The highest degree of income uncertainty across educational subgroups is found for the category "junior vocational". It contains people with a lower vocational education and those with an apprenticeship. Graduates from vocational colleges seem to have the lowest degree of uncertainty: there are relatively more people in this group who claim that their estimate of their future income is certain, and relatively less who are not very certain. Not displayed in the table are differences across evaluation horizon. For the 1 year horizon we find higher income uncertainty for the least educated. Our measure of income uncertainty does not lend support to the view that self–employed face higher uncertainty from income than employees: cf. the Pearson χ^2-statistic which indicates no significant differences. Cross–tabulating by financial wealth quintile gives the following picture: households with high income uncertainty are more concentrated in the lower end of the wealth distribution. Households in the upper quintile seem to have lower uncertainty.

4 Estimation

The subsequent analysis focuses on explaining the demand for riskless assets as a (budget) share in financial wealth. Consistent with an Engel curve of the Almost Ideal Demand System or the Translog specification, we consider the following equation:

$$y_{it}^* = x_{it}'\beta + \alpha_i + \epsilon_{it}; \quad i = 1, \ldots, N; \quad t = 1, \ldots, T. \tag{7}$$

where y^* is the desired share of riskless assets in financial wealth, x is a vector of explanatory variables, and ϵ denotes the error term. The inclusion of α_i denotes individual effects which have been ignored by other studies so far. We will come back to the importance of those effects later when we discuss panel data estimators. In cross–section analyses we have to ignore them. We treat y as a censored variable, implying that y^* is only exactly observed in the open interval $(0;1)$. Components of x are financial wealth, a measure of income uncertainty, and other variables which account for wealth accumulation and allocation (typically age and education or labor market variables).

This section discusses several ways of estimating the relation (7), both for cross–section and panel data, taking the censoring on y^* as far as possible into account. Presentation of results is deferred to section 5.

Cross–Section Estimators. An apparent first choice cross–section estimator for (7) is a parametric Tobit Maximum Likelihood (ML) estimator with a lower limit at 0 and an upper one at 1 under the assumptions of homoskedastic and normally distributed errors. However, if those assumptions are not met, the ML estimates will be inconsistent. One way to deal with heteroskedasticity is to allow the error variance to depend on covariates (for example, $\sigma_i = c \cdot \exp(z_i\gamma)$). Still, in the face of non–normality the parameter estimates might not be robust with respect to outliers.

The Tobit model assumes implicitly that the same underlying decision process drives both the participation decision and the decision of how much to invest in the asset under consideration. An alternative to this is the (standard) bivariate double hurdle model, which treats the censored observations, in particular the "1's" differently (cf. Blundell and Meghir (1987), for instance): a first hurdle — 'the participation equation' — determines in how far households are willing to hold non–riskless assets in the first place, whereas the censoring from the Tobit equation acts as the second hurdle which has to be overcome. Assumptions of normality and homoskedasticity are maintained for this estimator. We impose exclusion restrictions for identification. The double hurdle model is somewhat more appealing in terms of explaining economic outcomes than the Tobit, but due to its parametric assumptions it suffers from the same vulnerability with respect to misspecification as the Tobit. Therefore, we consider Tobit–type semiparametric estimators in the sequel.

Powell (1984) introduced as an alternative to the standard Tobit a conditional median estimator (CLAD) allowing for censoring. Powell (1986a) generalized the estimator to other quantiles. The model specifies a linear relationship between the latent variable y_i^* and regressors x_i, $y_i^* = x_i\beta_\theta + \epsilon_{\theta i}$ for the θ–th quantile. Only y_i is observed whose distribution is censored at yc_i ($yc_i = 1$ in our case). If censoring is from above, the estimator minimizes the objective function

$$\sum_{i=1}^{n} \rho_\theta(y_i - \min\{yc_i, x_i\beta_\theta\})$$

where $\rho_\theta(z)$ is the so–called "check function" $\rho_\theta(z) = (\theta - 1[z < 0])z$.

Errors are assumed to be independent, but not necessarily identically distributed. This allows for (bounded) heteroskedasticity. A conditional quantile restriction of the form $Q_\theta(\epsilon_{\theta i}|x_i) = 0$ is imposed. The conditional distribution function $F_{\epsilon_\theta}(\cdot|x)$ is assumed to be continuously differentiable around zero, and the conditional density is assumed to be positive at zero, $f_{\epsilon_\theta}(0|x) > 0$. Regressors are assumed to have finite support but need not be discrete. Inclusion of an intercept is required for consistency of the slope coefficients. In order to guarantee uniqueness of the estimator, the behavior of the regression function $x_i\beta_\theta$ has to be restricted. It is required that $x_i\beta_\theta \leq yc_i$ for sufficiently many observations. In addition, regressors must not be collinear for those observations. The censored quantile regression estimator is strongly consistent. Further regularity conditions are imposed to derive asymptotic normality.

The motivation for estimating quantile regressions is threefold: First, estimating several quantiles allows to get a more complete picture of the conditional distribution of $y|x$. Under homoskedasticity, the conditional quantile functions are parallel hyperplanes, but this need not hold under heteroskedasticity to which the estimator is fairly robust. Second, the estimator does not require specification of the error distribution as does the parametric Tobit. From a practical perspective, conditional median estimation (CLAD) would be the usual alternative to a standard Tobit, if one was only interested in the central tendency. But this proves infeasible with our data due to the high degree of censoring. We can obtain estimates for other quantiles below the median, though. Realize that when censoring is from above, the higher the quantile, the less precision can be expected from the estimator since then the quantile of y becomes uninformative about the true parameter value for large parts of the sample. This is the third reason to 'center' the distribution of y at lower quantiles. A short outline on optimization is given in Appendix C.

Standard errors can be obtained in different ways, cf. Appendix D. We use the Design Bootstrap Estimator which is based on re–sampling pairs of observations (cf. Efron and Tibshirani (1993)). Bootstrapping has the advantage that it avoids stringent assumptions on the underlying error distribution and the choice of smoothing parameters.

Panel Data Estimators. Consider estimating model (7) including individual effects α_i. A difference between cross–section and panel data models is that even in the absence of a dynamic specification the panel data estimator is more general in the following sense: fixed effects capture everything which is individually constant but unobserved. These might be related to preference parameters and indeed determine behavior, so that they have to be taken into account. They can be correlated with the regressors and even with the error terms.

Fixed effects estimators have an advantage over random effects estimators because the stochastic restrictions one has to place on the unobserved individual effects are less severe in the fixed effects case. In essence, they partly capture endogeneity of RHS variables. In addition, in the presence of attrition bias the estimates of the coefficients of interest, β, will be less seriously affected when a fixed effects estimator is used (Verbeek and Nijman

(1992)). A direct Tobit analog to the linear within estimator which differences fixed effects out does not exist under censoring. The parametric approach to estimate the FE Tobit model by ML would therefore have to estimate all individual effects. Estimates for α_i would be inconsistent as $T =$ fixed, which triggers inconsistency of the estimated β's.

Honoré (1992) devised a semiparametric Tobit type estimator for two waves of a panel, which takes account of fixed effects. The estimator comes in the form of a trimmed "least absolute deviations (LAD)" and a "least squares (LS)" variant. Honoré (1992) assumes that (y_1^*, y_2^*) are, conditional on (x_1, x_2, α), symmetrically distributed around the 45° degree line through the point $(\Delta x\beta, 0)$, where $\Delta x\beta = x_2 - x_1$. This follows directly if one assumes errors ϵ_1 and ϵ_2 between two waves to be continuously i.i.d. conditional on the regressors and fixed effects, (x_1, x_2, α). However, the estimator also allows for i.n.i.d. observations, and for correlation between error terms. Errors need not be homoskedastic across individuals, but they do have to be homoskedastic over time for a given individual. Both versions of the estimator are $\sqrt{N}$–consistent and asymptotically normal, yet asymptotically not efficient.[22] The idea is a generalization of Powell's (1986b) symmetrically censored least squares estimator for the cross–sectional case. As it leaves the error distribution unspecified, Honoré's estimator is thus far more general than feasible ML–based methods (i.e. the so–called "Chamberlain (1984)" approach).

In the application we will focus on the "LS" version of the estimator. The symmetry property under the above assumptions allows to derive a moment condition of the form[23]

$$\mathrm{E}(\chi(y_1, y_2, \Delta x\beta)|x_1, x_2) = 0 \tag{8}$$

where

$$\chi(z_1, z_2, \delta) = \begin{cases} z_1^2 + 2z_1(-z_2 - \delta) & \text{if} \quad \delta \leq -z_2 \\ (z_1 - z_2 - \delta)^2 & \text{if} \quad -z_2 < \delta < z_1 \\ z_2^2 + 2z_2(-z_1 + \delta) & \text{if} \quad z_1 \leq \delta \end{cases}$$

In order to establish the asymptotic properties and to estimate the covariance matrix, it is assumed that $\mathrm{E}(\Delta x' \Delta x)$ has full rank. The implication is that only time varying regressors can be used as conditioning variables. Time constant variables are subsumed under the fixed effects and therefore not estimated. The estimator minimizes the sample analog

$$T_n = \sum_{i=1}^{n} \chi(y_{i1}, y_{i2}, \Delta x_i \beta) \tag{9}$$

This objective function is globally convex in parameters and the minimum can be found using standard gradient methods.[24]

[22] Charlier et al. (1995) have devised an efficient GMM estimator based on Honore's (1992) FE Tobit estimator, involving non–parametric components. They report evidence that the efficiency gain can be small in small samples. Therefore, we do not resort to their estimator.

[23] In fact, Honore also derives moment conditions for both "LS" and "LAD" estimators under truncation. Those will be used in constructing a statistic to test for the symmetry assumption.

[24] The "LAD"–type estimator has a piecewise linear, convex objective function, which can be minimized

In order to apply Honoré's (1992) "LS" estimator to unbalanced panel data, we adopt the following procedure (cf. Charlier et al. (1995)): first, estimate the "LS" estimator $\beta_{\tau,\tau'}$ for every conceivable pair $\{(\tau,\tau')|1 \leq \tau < \tau' \leq T\}$ of panel waves. These estimates will differ from each other. Second, use a Minimum Distance (MD) step to restrict estimates to be the same, which results in a single estimator β_T. The MD estimator is calculated as

$$\beta_T = (\Psi'\Omega^{-1}\Psi)^{-1}\Psi'\Omega^{-1}\beta_H$$

where β_H are the first–round estimates, stacked in a vector, and Ψ is a matrix of linear restrictions between β_T and β_H. Ω is the covariance matrix of β_H used as weighting matrix in the MD step.[25] Although the resulting estimator is asymptotically not efficient, it is a consistent two–step estimator. The major advantages are that it allows for unbalanced panel data estimation and is straightforward to implement.

5 Portfolio responses to income uncertainty

In the special case where agents have CRRA utility, there is an indirect approach to determine whether precautionary motives matter for portfolio allocation, as should be clear from subsection 2.2. In a preliminary approach we estimated wealth elasticities separately for an array of assets. Inferring risk aversion from the estimated share equation, as is done here, is a more unifying approach, since the specification of the share is consistent with a general demand system. With both approaches, the estimates are consistent with DRRA. Since this does not allow to assess the existence of decreasing absolute prudence directly, some more identifying information is needed. This can come from our measures of income uncertainty and precautionary motives.

5.1 Specification and Findings

The main findings are summarized in tables 4–6 (Appendix B). In this section we discuss a baseline specification. A sensitivity analysis with respect to specification is provided in the following subsection. We will focus on the year 1993 when we discuss cross–section results.

Cross–Section Estimates. The baseline specification explains the share of riskless assets from age, education, financial wealth, labor market variables, family status, and

using simplex methods (cf. Press et al. (1989) for algorithms). It requires the choice of a smoothing parameter, though, for which little guidance is available. The latter seems to have prevented Honore (1992, section 6) from using it in his Monte Carlo–based performance check.

[25] Ω is calculated as $H^{-1}G'GH^{-1}$ where G is the matrix of first derivatives and H the (blockdiagonal) matrix of second derivatives of $\sum_{i=1}^{n}\sum_{\tau'=2}^{T}\sum_{\tau<\tau'}\chi(y_{i\tau}, y_{i\tau'}, \Delta_{\tau,\tau'}x_i\beta_{\tau,\tau'})$ wrt β_H.

income uncertainty. A flexible age pattern is allowed for by including a cubic polynomial. Education is included to proxy for life–time wealth (Hubbard et al. (1994), for example). The dummy variable for being married or cohabiting controls for family size. Financial wealth is modeled as a linear spline of log financial wealth. This serves to reduce the impact of measurement error while allowing for sufficient flexibility. As knots we chose points relating to the 20%–, 40%–, 60%–, 80%–percentiles of the distribution of financial wealth. The 'panel' variable distinguishes households from the so–called 'representative panel' (REP, the reference category) and the 'high–income panel' (HIP), and serves to control for income, which is not directly included in the equations due to severe problems with item non–response. From the budget–share approach it is wealth anyhow which is important, not income.

In the following discussion we focus on the effect of the income uncertainty on the investment in riskless, liquid assets. In the Tobit equation in table 4, the effect of income uncertainty has the expected sign: more income uncertainty leads to reduction in avoidable risks, here portfolio risk. Clearly, the higher the degree of income uncertainty the higher the share of financial wealth which the household invests in riskless assets. Thus, the found "effect" seems reasonable, even though the dummy variable for the category of moderate uncertainty is not significant. A Wald test does not reject the hypothesis that the coefficients are jointly zero.

Estimates can be affected by heteroskedasticity, even though the dependent variable is a share. A conditional moment test (cf. Newey (1985) and Pagan and Vella (1989)) suggests that heteroskedasticity is present (the test statistic has a value of 604.67 at 25 degrees of freedom (df); this value should be compared to the χ^2 distribution). Likewise, normality is rejected by a conditional moment test (127.42 at 2 df). Unlike the assumption of a particular family of the error distribution, heteroskedasticity can be dealt with in the Tobit by specifying the standard deviation of the error term to be dependent on other explanatory variables.[26] The $t-$values change somewhat, as is evident from table 4. An LR test supports the conclusion that heteroskedasticity matters in these data.

From the estimated coefficients we can compute wealth elasticities for the investment in risky assets. With the current specification we find:[27]

These estimates are consistent with decreasing relative risk aversion. But we cannot conclude that consumers have decreasing absolute prudence.

[26]The variance depends on age, the linear spline in log financial wealth, labor market variables and the dummy for panel membership (REP/HIP). This selection is arbitrary but parsimonious. Alternatives have been rejected on the basis of LR tests.

[27]The elasticity for risky assets can be derived from the estimates as $\eta = 1 - \beta_s/(1-y)$ where β_s is the spline s coefficient for log financial wealth in the share equation for riskless assets. Since this elasticity depends on the data, y, we calculate it at the mean, the first quartile and the median of y. Standard errors are obtained via the "delta" method.

spline of log financial wealth	mean of y		1st quartile of y		median of y	
1st quintile	3.07	(0.59)	2.03	(0.29)	6.80	(1.65)
2nd quintile	1.89	(0.28)	1.44	(0.14)	3.49	(0.80)
3rd quintile	1.71	(0.20)	1.36	(0.10)	3.00	(0.57)
4th quintile	1.77	(0.15)	1.39	(0.07)	3.17	(0.42)
5th quintile	1.53	(0.11)	1.26	(0.05)	2.47	(0.30)

Note: standard errors in parentheses.

Next to the Tobit estimates we present estimates for a double hurdle model, which treats the "1's" differently: the decision to hold only riskless, liquid assets may be driven by other considerations than the censoring story would imply. As exclusion restrictions we use dummy variables indicating very strong, strong, or general support for the following statement:

"I am very interested in financial matters (insurances, investments, etc.):" ....
. ... {answer on 7 point scale }

This set of questions has been asked to all respondents in the sample, irrespective of asset ownership. They can be interpreted as measure of informational status of the investor and should have a direct impact on the willingness to participate in markets for other than riskless and liquid assets.[28] The estimates are based on a bivariate specification, under the assumption of homoskedastic and normally distributed errors. The estimated correlation coefficient is significantly positive and with a value of 0.76 rather high. Comparing likelihood values between this model and the homoskedastic Tobit indicates that the introduction of the additional hurdle is a significant contribution to the model.[29] As apparent from the table, the effect of income uncertainty is slightly stronger, even though it is negligible in the participation equation. Judged by a Wald test, the coefficients of income uncertainty are jointly insignificant. Attempts to estimate the model under a heteroskedasticity specification (as done with the Tobit) failed.

Table 5a displays quantile regression models which take account of censoring. The covariance matrix was estimated by bootstrapping pairs of observations (y_i, x_i). The bootstrap samples were of the same size as the original sample. 250 bootstrap iterations

[28] This is the hypothesis maintained by King and Leape (1987) who see information as a core determinant of portfolio decisions.

[29] Tobit and double hurdle are nested if one restricts the probability of the participation equation to be unity, cf. Blundell and Meghir (1987). Note that our specification is somewhat more general than the one discussed in Blundell and Meghir (1987) as we do not restrict the correlation coefficient between the equations to be zero.

were used.[30] This estimator is robust to heteroskedasticity.

Compared to the censored mean regressions (Tobits), the parameter estimates change somewhat in terms of size and significance levels. The indicator for high income uncertainty has a large significant impact on the share invested in riskless wealth only for the 0.2 quantile. Both estimates and $t-$values first increase and then decrease as the quantiles increase. A Wald test does again not reject the hypothesis that both dummies measuring income uncertainty have zero coefficient.

Again the estimates suggest that as financial wealth goes up, the share of riskless wealth goes down (DRRA). As we go across quantiles, the effect of income uncertainty on holding riskless assets is non–linear when the share of riskless wealth increases. Households in the 0.2 quantile of the share–of–riskless–assets distribution are more sensitive to income uncertainty. Clearly, they hold more risky portfolios than those in higher quantiles. Beyond the 0.2 quantile, the effect of income uncertainty is reduced (both in magnitude and in significance). Statistically, the differences across quantiles are an indication for heteroskedasticity; in a homoskedastic sample the conditional quantile estimates should follow the central tendency of $y|x$. Buchinsky (1996) outlines a Minimum Distance approach to test for heteroskedasticity. Under the Null of homoskedastic errors, the devised test statistic is based on the equality of slope parameters across different quantiles. The test statistic takes a value of 137.83 at 72 degrees of freedom. Compared to the percentiles of a χ^2- distribution, homoskedasticity is rejected at the 5% level.[31]

The quantile estimator also allows to trace out the entire conditional distribution of the dependent variable by varying the parameter θ from 0 to 1. This is displayed in figure 1. We have estimated all percentiles of the distribution of $y|x$ and averaged the predicted percentile $\hat{Q}_\theta(y_i|x_i) = \min\{x_i\hat{\beta}_\theta, yc_i\}$ over all individuals i by subsample, which results in a picture of the empirical distribution function, conditional on income uncertainty.[32] Again it is apparent that the higher income uncertainty, the more probability mass is concentrated around high shares (indicating riskless portfolios).

Panel Data Estimates. Table 6a displays results from applying the Honoré (1992) estimator. Censoring is from above at 1. The variables chosen are the wealth variables and income uncertainty. Other variables present in the cross–section specification do not appear explicitly since they are fixed over time for a given individual (like education).[33] They are taken care of by the fixed effects. Starting values have been obtained from the standard linear within estimator, ignoring censoring. The first three sets of columns

[30] Note that bootstrapping residuals is not applicable since residuals are unobservable due to censoring. Cf. Appendix D for further details.

[31] For this test, Powell's (1986a) estimator of the covariance matrix under homoskedasticity was used, bandwidth selection was guided by the Hall and Horowitz (1990) method; cf. Appendix D.

[32] Due to non–convergence at higher quantiles, we can only display a part of the function.

[33] Many other variables did not have a sufficient within–variation. For labor market variables the information has not been requested in the 1994 wave.

display the Honoré estimates for all pairs of available waves (pairwise balanced), the last two columns combines those estimates into a single one using a MD step (unbalanced).

The financial wealth spline is still significant, as it was in the cross sectional models. We have chosen the same kink points for the splines across waves, i.e. we fixed the knots at the 1993 quintiles for reason of comparability. Also, financial wealth has been deflated using the consumer price index. The effect of income uncertainty is not as clear as it was in the cross sectional Tobit specifications, with the exception of the non–adjacent waves 1993 and 1995. The effect vanishes completely for the MD estimates. This might entirely be due to the individual fixed effects which the panel data estimators take into account. Apparently, there is a positive correlation between fixed effects and income uncertainty: once fixed effects are taken into account, they take away some of the explanatory power of income uncertainty. Clearly, if we think of the estimator of a within–type which differences all variables from their individual means over time, then the "permanent" component of income uncertainty is differenced out. We are left with "temporary" uncertainty which we should not expect to have any explanatory power for savings or portfolio choice. But note as well that the estimator is not efficient. Moreover, the underlying assumptions of the estimator may not be met by the data. This concerns in particular the symmetry assumption. A conditional moment test rejects symmetry at the 5% level (test statistic: 79.80, 6 df).[34]

5.2 Sensitivity Analysis

This section considers deviations from the baseline specification into several directions, both for the cross–sectional censored quantile regressions and for the panel data estimates.

In the cross–sectional case we first consider the 'income uncertainty' measure for a horizon of 12 months (panel b of the tables presented below). We then look at the alternative measures (c) 'income variance' and (d and e) 'precautionary motives'. Subsequently, we explore the sensitivity of the specification wrt to some RHS variables, taking account of interaction terms. In the baseline specification we do not take into account the level of the income expectation. However, due to the theoretical requirement that the marginal utility be convex, it is conceivable that the effects of income uncertainty on the share of riskless assets will depend on the level of the expected income change (f). Income expectations are measured by subjective point estimates. Further, we use interaction effects of income uncertainty with wealth (g). Next, if financial wealth is endogenous, is the specification sensitive to either dropping those variables altogether (this would be a "reduced form" interpretation of the estimated equation, h) or to using predicted value as a continuous measure on the RHS (i)? We also explore whether pension entitlements

[34] The moments for this test are based on (8) and the moments of the truncated version of the "LS" estimator which is not used in estimation. Under symmetry both those moment conditions have to be met. Cf. Honore (1992) and Charlier et al. (1995).

(j) should be used as a conditioning variable, and address the impact of bequest motives as an alternative saving motive (k). Also, the level of inflation expectation a predictor of the asset allocation (l) will be considered.

Cross–Section Estimates. Table 5b displays the coefficients of variables which are directly of interest. The subpanels are numbered as explained above. The difference in the number of observations is due to item non–response in at least one of the underlying variables.

Interestingly, in panel b the short–run measure for high income uncertainty is significant in some of the quantile regressions. From the theoretical discussion above, we would have expected the long–run measure to be a more appropriate measure. The presumption that transaction costs would be of particular relevance in cross–section estimates, which should be reflected in a stronger effect for the long–run measure, does not seem to be supported. Another possibility is that households are myopic and do not consider risks which are far away in time for portfolio allocation.

The continuous measure of income variance (c), which is also based on a 1 year horizon, yields hardly significant estimates, though.

The measure for 'precautionary motives' shows a significant parameter estimate for the 0.2 and 0.3 quantiles in the broad interpretation (d). This measure captures other related risks as well, like health risks, unemployment risks, and risks from outstanding debts (the latter could be interpreted as future liquidity risks). The narrow measure (unforeseen expenditure alone) (panel e) is insignificant throughout.

In panels f and g we interact income uncertainty with income expectation and wealth, respectively. Wald tests suggest that the interaction terms are jointly insignificant.

Financial wealth might be endogenous. One reason to suppose this is that if income uncertainty is a life–cycle phenomenon, households will respond with a higher level of saving to increased background risk. These additional savings are channeled into all different sorts of assets and thus increase financial wealth as such. There are two short–cuts to deal with the problem: either one excludes financial wealth altogether, hoping that the other variables in the specification will pick up its effect (reduced form), or one predicts log financial wealth from a set of other variables. We use predictions from a censored median regression (with individual censoring points) of the log of financial wealth on most of the other socio–demographic characteristics displayed in the other tables, plus the variables on financial interest, cf. p. 20 , which serve as exclusion restrictions.[35]. This is done in the subsequent panels (h and i) in the table. Although changes are observed for some of the point estimates, the conclusions are unaffected. We would therefore not view financial wealth to cause a considerable endogeneity bias for income uncertainty.[36]

[35]Estimates are available upon request.

[36]Lee (1995) proposes a way to deal with endogeneity in the following way: first estimate reduced form equations for both the share equation and an equation explaining financial wealth; then, impose

Taking account of pension entitlement, bequest motives, or price expectations has only negligible impacts on the coefficients of income uncertainty (cf. panels j through l). This can be interpreted as weak evidence that neither the existence of old–age income provision, nor competing saving motives, nor inflation bias the findings reported for the baseline case.

Panel Data Estimates. A sensitivity analysis following the same lines is provided in table 6b. The table displays MD estimates, except for the cases where variables of interest are not available for one of the waves. As in the cross–sectional case we conclude that the baseline specification is robust to the changes considered.

5.3 Closing remark on estimates

What have we learned from the data? The estimates presented in this paper lend some but not unanimous support to the hypothesis that uninsurable earnings risk as perceived by the individual household leads to reduction in avoidable portfolio risk. Especially with panel data models the hypothesis of income uncertainty not causing any significant response in household portfolios is hard to reject. As is frequently noted in the applied literature, estimates lose precision when ML is abandoned. This might be an explanation for the lower $t-$values we find for the semiparametric estimates. Moreover, the Honoré estimator in particular is not efficient, and this will have some bearing on the estimated standard errors as well. On the other hand, as there is evidence for heteroskedasticity of unknown form, we prefer to employ the robust methods.

6 Conclusions

We have empirically investigated the question whether income uncertainty matters for portfolio allocation of households. To this end, we employed a new data source from the Netherlands, and estimated budget share equations for liquid and riskless financial asset in financial wealth. Emphasis was put on using robust econometric methods. This point is important since theory gives hardly any guidance as how to evaluate the question at hand econometrically. In particular, we contrast parametric estimators with semiparametric methods which avoid the specification of an error distribution, or place less restrictions on it than ML would. Censoring is taken into account in all specifications. In panel data estimates we allow for fixed effects.

parameter restrictions between equations by minimum distance to obtain structural form estimates. This approach is precluded here, since the latter step involves the choice of an appropriate bandwidth, for which we have not found a good method, cf. Appendix D.

As indicator for income uncertainty we relied upon a self–reported measure with categorical information. Our measure has two main advantages: first, since households have more information about their income development than any outside observer, we do not need to base the evaluation of future earnings risk on the observed earnings history of the household; this also avoids misspecification of the estimated model, since a continuous measure of some expected income variance implies again certain restrictions on the equation of interest (see the criticism in Carroll and Samwick (1995b)). Second, with a 5 year horizon it is of a considerably longer term than subjective measures employed elsewhere. We can also assess the differential effects of using a shorter horizon.

Our findings from different kinds of estimators are somewhat inconclusive. Viewing the range of predictions for the level of precautionary wealth as reported in the existing literature (again, see the overview in Carroll and Samwick (1995b)) this is hardly surprising. If anything, our estimates suggest that income uncertainty may indeed be a source for precautionary saving and also induces more riskless portfolios. But the effect is not always clear–cut.

There are several ways of interpreting these findings. Empirically, the fixed effects model might be capturing temporary components of uncertainty and difference out permanent uncertainty.

From a theoretical perspective, one might conclude that preferences do not meet the requirements of decreasing absolute prudence (we did find evidence for decreasing relative, and hence, absolute risk aversion). Note in particular, that standardness is a restrictive assumption, compared with other concepts (like risk vulnerability, Gollier and Pratt (1996)).

Comparing with the findings from other countries (for the US (Chakraborty and Kazarosian (1996)) and Italy (Guiso et al. (1996)), one should realize that the Netherlands have a quite well organized system of social security provisions and a huge sector of occupational pensions which would to some extent reduce the need for precautionary saving, at least for the risk from earnings.

References

Allen, F. and D. Gale (1994): *Financial Innovation and Risk Sharing.* Cambridge, MA: MIT Press.

Andrews, D. W. K. and M. Buchinsky (1996): "On the Number of Bootstrap Repetitions for Bootstrap Standard Error Estimates," Cowles Foundation, mimeo.

Barrodale I. and F. D. K. Roberts (1973): "An Improved Algorithm for Discrete ℓ_1 Linear Approximation," *SIAM Journal of Numerical Analysis,* 10, 839–848.

Bertaut, C. C. and M. Haliassos (1992): "Precautionary Portfolio Behavior," Working Paper 92-8, University of Maryland at College Park.

Bloomfield and Steiger (1983): *Least Absolute Deviations. Theory, Applications, and Algorithms.* Boston: Birkhäuser.

Blundell, R. and C. Meghir (1987): “Bivariate Alternatives to the Tobit Model,” *Journal of Econometrics,* 34, 179-200.

Blundell, R. and T. Stoker (1994): “Consumption and the Timing of Income Risk,” IFS Working Paper, W94/9.

Browning, M. and A. Lusardi (1996): “Household Saving: Micro Theories and Micro Facts,” *Journal of Economic Literature,* 34, 1797–1855.

Buchinsky, M. (1994): “Changes in the U.S. Wage Structure 1963–1987: Application of Quantile Regression,” *Econometrica,* 62, 405–458.

—— (1996): “Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research,” Yale University, mimeo.

Camphuis, H. and M. Ketelaars (1995): “Variable Documentation VSB Panel 1993,” Tilburg University.

Carroll, C. D., and A. A. Samwick (1995a): “The Nature of Precautionary Wealth,” NBER Working Paper, No. 5193.

—— (1995b): “How Important is Precautionary Saving?” NBER Working Paper, No. 5194.

Chakraborty, A. and M. Kazarosian (1996): “Portfolio Allocation of Precautionary Assets: U.S. Panel Evidence,” Brandeis University and Northeastern University, mimeo.

Chamberlain, G. (1984): “Panel Data,” in *Handbook of Econometrics,* ed. by Z. Griliches and M. Intriligator. Amsterdam: North–Holland, 1247–1318.

—— (1994): “Quantile Regression, Censoring, and the Structure of Wages,” in *Proceedings of the 6th World Congress of the Econometric Society, Barcelona,* ed. by C. Sims and J. J. Laffont. New York: Cambridge University Press.

Charlier, E., B. Melenberg, and A. van Soest (1995): “Estimation of a Censored Regression Panel Data Model Using Conditional Moment Restrictions Efficiently.” CentER DP No. 95114, Tilburg University.

Das, M. and A. van Soest (1996): “A Panel Data Model for Subjective Information on Household Income Growth,” VSB–CentER Savings Project, Progress Report No. 41, Tilburg University.

Davies, J. B. (1981): “Uncertain Lifetime, Consumption, and Dissaving in Retirement,” *Journal of Political Economy,* 89, 561–577.

Deaton, A. (1992): *Understanding Consumption.* Oxford: Clarendon Press.

Drèze, J. H. and F. Modigliani (1972): “Consumption Decisions under Uncertainty,” *Journal of Economic Theory,* 5, 308–335.

Eeckhoudt, L. and C. Gollier (1995): *Risk: Evaluation, Management and Sharing.* New York et al.: Harvester Wheatsheaf.

Efron B. and R. J. Tibshirani (1993): *An Introduction to the Bootstrap.* New York and London: Chapman and Hall.

Elmendorf, D. W. and M. S. Kimball (1991): “Taxation of Labor Income and the Demand for Risky Assets,” NBER Working Paper No. 3904.

Fitzenberger, B. (1994): ”A Note on Estimating Censored Quantile Regressions,” Center of Int’l. Labor Economics, DP 14, University of Konstanz.

—— (1997): “A Guide to Censored Quantile Regressions,” In: G.S. Maddala and C.R. Rao (eds.), Vol. 15 of the *Handbook of Statistics: Robust Inference*. Amsterdam: North–Holland, (*forthcoming*).

Gollier, C. and J. W. Pratt (1996): “Risk Vulnerability and the Tempering Effect of Background Risk,” *Econometrica,* 64, 1109–1123.

Guiso, L., T. Jappelli and D. Terlizzese (1992): “Earnings Uncertainty and Precautionary Saving,” *Journal of Monetary Economics,* 30, 307–337.

—— (1996): “Income Risk, Borrowing Constraints and Portfolio Choice,” *American Economic Review,* 86, 158–172.

HAHN, J. (1995): “Bootstrapping Quantile Regression Estimators,” *Econometric Theory,* 11, 105–121.

HALL, P. AND J. HOROWITZ (1990): “Bandwidth Selection in Semiparametric Estimation of Censored Linear Regression Models,” *Econometric Theory,* 6, 123–150.

HENDERSON, J. V. AND Y. M. IOANNIDES (1983): “A Model of Housing Tenure Choice,” *American Economic Review,* 73, 98–113.

HOCHGUERTEL, S., R. ALESSIE, AND A. VAN SOEST (1997): “Saving Accounts versus Stocks and Bonds in Household Portfolio Allocation,” *Scandinavian Journal of Economics,* 99, 81–97.

HOCHGUERTEL, S. AND A. VAN SOEST (1996): “The Relation between Financial and Housing Wealth of Dutch Households,” CentER DP No. 9682, Tilburg University. `http://greywww.kub.nl:2080/greyfiles/center/1996/82.html`

HONORÉ, B. E. (1992): “Trimmed LAD and Least Squares Estimation of Truncated and Censored Regression Models with Fixed Effects,” *Econometrica,* 60, 533–565.

HOROWITZ, J. (1997): ”Bootstrap Methods for Median Regression,” University of Iowa, mimeo.

HUBBARD, G. R., J. SKINNER AND S. P. ZELDES (1994): “The Importance of Precautionary Motives in Explaining Individual and Aggregate Saving,” *Carnegie Rochester Conference Series on Public Policy*, 40, 59–126.

HUBER, P. J. (1981): *Robust Statistics.* New York et al.: John Wiley & Sons.

JIANAKOPLOS, N. A., P. L. MENCHIK, AND F. O. IRVINE (1996): “Saving Behavior of Older Households: Rate–of–return, Precautionary and Inheritance Effects,” *Economics Letters,* 50, 111–120.

KAZAROSIAN, M. (1997): “Precautionary Savings — A Panel Study,” *Review of Economics and Statistics, forthcoming.*

KIMBALL, M. S. (1990): “Precautionary Saving in the Small and in the Large,” *Econometrica,* 58, 53–73.

—— (1992): “Precautionary Motives for Holding Assets,” in P. Newman, M. Milgate and J. Eatwell (eds.), *The New Palgrave Dictionary of Money and Finance,* vol. III. London: Macmillan, pp. 158–161.

—— (1993): “Standard Risk Aversion,” *Econometrica,* 61, 589–611.

KING, M. A. AND J. I. LEAPE (1987): “Asset Accumulation, Information, and the Life Cycle,” NBER Working Paper, No. 2392.

KOENKER, R. AND G. BASSETT (1978): “Regression Quantiles,” *Econometrica,* 46, 33–50.

KOENKER, R. AND B. J. PARK (1996): “An Interior Point Algorithm for Nonlinear Quantile Regression,” *Journal of Econometrics,* 71, 265–283.

KOO, H. K. (1991): “Consumption and Portfolio Choice with Uninsurable Income Risk,” Princeton University, mimeo.

LEE, M. J. (1995): “Semi–Parametric Estimation of Simultaneous Equations with Limited Dependent Variables: A Case Study of Female Labor Supply,” *Journal of Applied Econometrics,* 10, 187–200.

LELAND, H. A. (1968): “Saving and Uncertainty: The Precautionary Demand for Saving,” *Quarterly Journal of Economics,* 465–473.

LUSARDI, A. (1996): “Precautionary Saving and Subjective Earnings Variance,” Dartmouth College and ICER, mimeo.

NEWEY, W. (1985): “Maximum Likelihood Specification Testing and Conditional Moment Tests,” *Econometrica*, 53, 1047–1070.

PAGAN, A. AND F. VELLA (1989): “Diagnostic Tests for Models Based on Individual Data,” *Journal of Applied Econometrics,* 4, S29–S59.

POWELL, J. (1984): “Least Absolute Deviations Estimation for the Censored Regression Model,” *Journal of Econometrics,* 25, 303–325.

—— (1986a): "Censored Regression Quantiles," *Journal of Econometrics,* 32, 143–155.

—— (1986b): "Symmetrically Trimmed Least Squares Estimation for Tobit Models," *Econometrica,* 54, 1435–1460.

Pratt, J. W. and R. Zeckhauser (1987): "Proper Risk Aversion," *Econometrica,* 55, 143–154.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling (1989): *Numerical Recipes: The Art of Scientific Computing.* Cambridge et al.: Cambridge University Press.

Rothschild, M. and J. Stiglitz (1970): "Increasing Risk: I. A Definition," *Journal of Economic Theory,* 2, 225–243.

Verbeek, M. and T. Nijman (1992): "Testing for Selectivity Bias in Panel Data Models," *International Economic Review,* 33, 681–703.

Womersley, R. S. (1986): "Censored Discrete Linear ℓ_1 Approximation," *SIAM Journal of Scientific Statistical Computing,* 7, 105–122.

Zeldes, S. P. (1989): "Optimal Consumption with Stochastic Income: Deviations from Certainty Equivalence," *Quarterly Journal of Economics*, 104, 275–298.

Appendix A: Measures of Income Risk

Here we list the information provided in the data[37].

(*a*)

[1899.] *"Do you think, taking into account possible changes within the household, the total net income of your household will increase, remain the same, or decrease, <u>in the next 12 months</u>?":* . . . . **1 -** increase, **2 -** remain the same, **3 -** decrease.

conditional upon answers **1** or **3**:

(*b*1)

[1900.] *"By what <u>percentage</u> do you think the total net income of your household will increase <u>in the next 12 months</u>?":* . . . . in **%-age** points.

(*b*3)

[1901.] *"By what <u>percentage</u> do you think the total net income of your household will decrease <u>in the next 12 months</u>?":* . . . . in **%-age** points.

(*c*)

[1902.] *"How certain do you feel about this change of income?":* . . . . **1 -** very certain, **2 -** rather certain, **3 -** not very certain **4 -** not at all certain.

$(d), 1-7$

[1903–1909.] *Consider a change of your income by xx% next year: how likely do you think is this?:*[38] . . . . **1 -** highly unlikely, . . ., **7 -** highly likely.

There are seven questions, one each for *xx%*: $< -15\%$; $-15/-10\%$; $-10/-5\%$; $-5/+5\%$; $+5/+10\%$; $+10/+15\%$; $> +15\%$.

(*e*)

[1910.] *"Do you think the total net income of your household will increase, remain the same, or decrease, in the <u>next five years</u>?":* . . . . **1 -** increase, **2 -** remain about the same, **3 -** decrease.

conditional upon answers **1** or **3**:

(*f*1)

[37] cf. Camphuis and Ketelaars (1995); variable numbers refer to numbers therein; 1993 wave.

[38] the wording in the questionnaire is different, although the content is the same as here.

[1911.] *"By what percentage do you think the total net income of your household will increase in the next five years?":* .... in **%-age** points.
$(f3)$
[1912.] *"By what percentage do you think the total net income of your household will decrease in the next five years?":* .... in **%-age** points.
(g)
[1913.] *"How certain do you feel about this change of income?":* ....**1 -** very certain, **2 -** rather certain, **3 -** not very certain **4 -** not at all certain.
(h)
[1914.] *"Do you expect prices in general to rise, to remain the same, or to go down, in the next 12 months?":* .... **1 -** go down, **2 -** remain the same, **3 -** rise.
conditional upon answer 3:
(i)
[1915.] *"By how many percent do you expect prices to rise in the next 12 months?":* ... . in **%-age** points.
(j)
[1916.] *"Do you expect prices in general to rise, to remain the same, or to go down, in the next five years?":* .... **1 -** go down, **2 -** remain the same, **3 -** rise.
conditional upon answer 3:
(k)
[1917.] *"By how many percent do you expect prices to have risen after five years?":* .... in **%-age** points.

The measure 'income uncertainty' is a set of two dummy variables (three categories), directly obtained from variables (c) and (g), for the 1 and 5 year horizon, respectively: category 1 is the reference group, category 2 is labeled 'moderate uncertainty' and categories 3 and 4 are dubbed 'high uncertainty'.

To obtain a measure of income uncertainty which is closer related to a measure of variance, the following procedure has been suggested by Rob Alessie (source: personal communication):

Rescale the variables (d) which ask for the likelihood of a given income change such that the answer categories (1-7) match into the interval (0.05,0.95) and interpret these answers as probabilities. Then form an expected value by dividing the new variables by the sum over all 7 questions (ie. the possible given ranges for the income change). From there, construct the variance for 1–year income uncertainty. A technical problem is the arbitrary assumption that the answer categories are evenly spaced (ie. the probability difference between 'likely' and 'not so likely' is the same as between 'highly unlikely' and 'unlikely').

Appendix B: Tables and Figures

Figure 1:

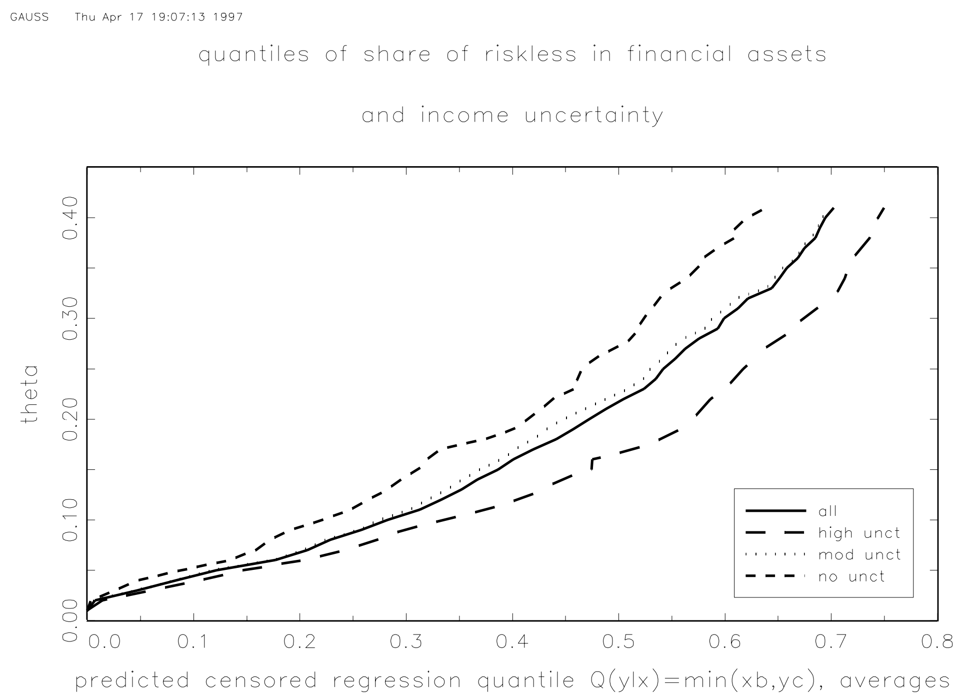

Table 1: Risk Aversion and the Demand for Risky Assets

	DARA	CARA	IARA
DRRA	$\eta > 1$	./.	./.
CRRA	$\eta = 1$	./.	./.
IRRA	$0 < \eta < 1$	$\eta = 0$	$\eta < 0$

η: elasticity of the demand
for risky assets wrt initial wealth
./.: inconsistent with expected utility theory

table adapted from Eeckhoudt and Gollier (1995), Table 9.1, p. 134

Table 2: Income Expectation and Income Uncertainty
— heads of households, in % —

a) 1 year horizon, 1993 (nobs. = 1915)

expect income to	%-age of sample	income uncertainty low	moderate	high	all
increase	16.40	26.43	57.64	15.92	100.00
same	70.18	24.40	66.29	9.30	100.00
decrease	13.42	26.85	49.42	23.74	100.00
all	100.00	25.07	62.61	12.32	100.00

Pearson $\chi^2_4 = 52.52$; $p-$value $= 0.0\%$

b) 5 year horizon, 1993 (nobs. = 1915)

expect income to	%-age of sample	income uncertainty low	moderate	high	all
increase	35.40	9.73	64.90	25.37	100.00
same	46.06	12.59	69.16	18.25	100.00
decrease	18.54	15.21	55.21	29.58	100.00
all	100.00	12.06	65.07	22.87	100.00

Pearson $\chi^2_4 = 30.79$; $p-$value $= 0.0\%$

c) 1 year against 5 year horizon, 1993 (nobs. = 1915)

1 year/5 yrs	low	moderate	high	total
low	8.88	14.36	1.83	25.07
moderate	3.08	46.89	12.64	62.61
high	0.10	3.81	8.41	12.32
all	12.06	65.07	22.87	100.00

Pearson $\chi^2_4 = 623.79$; $p-$value $= 0.0\%$

d1) 5 year horizon, 1993 against 1994 (nobs. = 1486)

1993 /1994	low	moderate	high	total
low	4.98	5.99	0.54	11.51
moderate	6.80	46.57	12.38	65.75
high	0.74	10.36	11.64	22.75
all	12.52	62.92	24.56	100.00

Pearson $\chi^2_4 = 323.33$; $p-$value $= 0.0\%$

d2) 5 year horizon, 1994 against 1995 (nobs. = 1258)

1994 /1995	low	moderate	high	total
low	6.04	6.60	1.19	13.83
moderate	7.00	45.87	8.98	61.84
high	0.79	13.04	10.49	24.32
all	13.83	65.50	20.67	100.00

Pearson $\chi^2_4 = 265.46$; $p-$value $= 0.0\%$

d3) 5 year horizon, 1993 against 1995 (nobs. = 1042)

1993 /1995	low	moderate	high	total
low	4.99	5.85	0.67	11.52
moderate	6.81	48.37	10.84	66.03
high	1.34	11.90	9.21	22.46
all	13.15	66.12	20.73	100.00

Pearson $\chi^2_4 = 177.62$; $p-$value $= 0.0\%$

Table 3: background variables, share of riskless assets, and income uncertainty
— heads of households, in %, 1993 —

Nobs = 1915	% of sample	share of riskless assets: mean	pc10[a]	pc20	pc30	pc40	income uncertainty: low	moderate	high
age class									
under 25	2.6	.90	.63	.87	1	1	12.0	50.0	38.0
25–39	34.3	.74	.14	.40	.63	.78	8.7	65.2	26.1
40–54	36.8	.64	.08	.23	.38	.53	11.5	65.6	22.9
55–64	14.2	.67	.09	.24	.45	.58	18.8	68.8	12.5
65+	12.2	.74	.09	.33	.64	.85	15.5	61.8	22.7
							Pearson $\chi^2_8 = 42.13$; p–value = 0.0%		
education level									
primary education	5.4	.81	.19	.55	.82	1	9.7	65.0	25.2
secondary educ.	9.9	.76	.12	.33	.73	.94	7.9	68.4	23.7
pre–university	12.5	.69	.11	.24	.47	.66	13.0	64.9	22.2
junior vocational	11.3	.75	.09	.39	.65	.97	7.9	61.6	30.6
senior vocational	10.4	.68	.10	.26	.47	.68	12.6	67.8	19.6
vocat. college	29.7	.70	.13	.32	.52	.71	14.4	66.2	19.4
university	20.9	.62	.06	.21	.36	.53	12.8	62.5	24.8
							Pearson $\chi^2_{12} = 21.39$; p–value = 4.5%		
employment status									
self–employed	9.9	.68	.03	.18	.41	.68	11.6	63.5	24.9
unemployed	1.7	.91	.52	.96	1	1	3.1	62.5	34.4
the rest	88.5	.70	.11	.29	.50	.69	12.3	65.3	22.4
							Pearson $\chi^2_4 = 4.65$; p–value = 32.5%		
panel									
representative	61.5	.77	.14	.46	.69	.92	11.4	63.1	25.5
high income	38.5	.58	.07	.19	.30	.47	13.1	68.2	18.7
							Pearson $\chi^2_2 = 12.07$; p–value = 0.2%		
financial wealth									
1st quintile	20.0	.91	.68	1	1	1	10.7	59.3	30.0
2nd quintile	20.0	.81	.33	.54	.75	.94	11.7	64.8	23.5
3rd quintile	20.0	.76	.23	.48	.65	.82	10.4	67.1	22.5
4th quintile	20.0	.60	.09	.25	.35	.50	11.7	66.6	21.7
5th quintile	20.0	.41	.02	.07	.13	.21	15.7	67.6	16.7
							Pearson $\chi^2_8 = 23.61$; p–value = 0.3%		
income uncertainty									
low	12.1	.65	.07	.23	.38	.55			
moderate	65.1	.69	.10	.28	.50	.68			
high	22.9	.75	.15	.40	.64	.82			
Total	100.0	.70	.11	.29	.50	.70	12.1	65.1	22.9

Note: figures for 1994 and 1995 available upon request
[a] pc10 (etc.) means: 10th (etc.) percentile

Table 4: Parametric Estimates
— Baseline Specification, 1993 —
explained: share of riskless in financial wealth

	homoskedastic Tobit		**heteroskedastic* Tobit**		**Bivariate Double Hurdle**** **Tobit equation**		**particip. eqn.**	
	coeff.	t-value	coeff.	t-value	coeff.	t-value	coeff.	t-value
constant	3.4907	(6.939)	6.9555	(4.46)	1.6586	(2.805)	–3.5645	(–1.645)
age	–0.0900	(–3.154)	–0.0159	(–0.459)	–0.1088	(–3.116)	0.1453	(1.091)
$age^2/100$	0.1780	(3.012)	0.0044	(0.061)	0.2150	(2.980)	–0.3437	(–1.257)
$age^3/1000$	–0.0107	(–2.761)	0.0021	(0.425)	–0.0132	(–2.809)	0.0242	(1.367)
secondary educ.	–0.0254	(–0.366)	–0.0375	(–0.490)	–0.0829	(–1.030)	- 0.0158	(–0.053)
pre–university	–0.0967	(–1.454)	–0.1135	(–1.531)	–0.1171	(–1.474)	0.1293	(0.437)
junior vocat.	–0.0930	(–1.359)	–0.0769	(–0.998)	–0.1440	(–1.766)	0.1799	(0.624)
senior vocat.	–0.1806	(–2.653)	–0.1715	(–2.237)	–0.1746	(–2.200)	0.4488	(1.485)
vocat. college	–0.0437	(–0.708)	–0.0434	(–0.615)	–0.0570	(–0.766)	0.1943	(0.697)
university	–0.0763	(–1.195)	–0.0599	(–0.832)	–0.0886	(–1.145)	0.2946	(0.959)
diploma	0.0662	(2.009)	0.0789	(2.277)	0.0671	(1.898)	–0.0487	(–0.363)
log fin. wealth 1	–0.1049	(–3.263)	–0.6227	(–3.507)	0.1176	(3.469)	0.1686	(2.027)
log fin. wealth 2	–0.1935	(–3.892)	–0.2676	(–3.121)	–0.0728	(–1.254)	0.2978	(1.615)
log fin. wealth 3	–0.1789	(–3.117)	–0.2143	(–3.512)	–0.1266	(–1.968)	0.3595	(1.377)
log fin. wealth 4	–0.2612	(–5.365)	–0.2330	(–5.199)	–0.2180	(–3.988)	0.8780	(3.012)
log fin. wealth 5	–0.1397	(–4.217)	–0.1581	(–4.871)	–0.1287	(–3.166)	–0.0656	(–0.287)
self–employed	0.0237	(0.542)	0.0484	(0.969)	–0.0219	(–0.491)	–0.1393	(–0.802)
no paid job	0.0878	(1.095)	0.0514	(0.671)	–0.0066	(–0.071)	–0.1647	(–0.511)
other job	0.0435	(0.571)	0.0554	(0.686)	–0.0120	(–0.146)	–0.1751	(–0.500)
retired / disabled	–0.0290	(–0.604)	–0.0429	(–0.939)	–0.0453	(–0.838)	–0.0062	–0.028)
unemployed	0.3275	(2.638)	0.2257	(2.006)	0.1995	(0.800)	–0.6765	(–0.882)
married or cohab.	–0.0686	(–2.037)	–0.0570	(–1.627)	–0.0079	(–0.214)	0.2321	(1.687)
panel (1=HIP)	–0.0842	(–2.801)	–0.1003	(–3.423)	–0.0598	(–1.771)	0.3857	(2.444)
high income unc.	0.0869	(2.010)	0.0877	(2.037)	0.0896	(1.820)	–0.1991	(–1.010)
moder. inc. unc.	0.0514	(1.388)	0.0618	(1.680)	0.0633	(1.507)	–0.0895	(–0.512)
financ. int.: 5					—	(—)	0.5170	(3.151)
financ. int.: 6					—	(—)	0.6384	(2.869)
financ. int.: 7					—	(—)	0.5049	(2.093)
nobs	1915		1915		1915			
left cens. at 0	43		43		43			
right cens. at 1	823		823		823			
Log Lhd.	–1242.2		–1159.8		–1034.9			
Wald test	4.060	(2 df)	4.230	(2 df)	3.394	(2df)	1.169	(2df)
Wald test						6.920	(4df)	
Pseudo R^2	0.2288		0.2800		0.3604			

Definition of variables: highest level of education: lowest education (primary/elementary or special low level education; reference cat.); secondary educ.: continued special (low level) education, secondary educ.; pre–university; junior vocat.: junior vocational education and apprenticeships; senior vocat.: senior vocational training; vocat. college: vocational colleges; university; diploma: received degree; log fin. wealth: linear spline with 4 knots at the quintiles of the distribution of log(fin. wealth+1); self–employed through unemployed: reference cat.: paid job. uncertainty: cf. text and Appendix A. risk attitude: prepared to take risks, 7–pt scale (cf. text); financ. int.: interested in financial matters, 7–pt scale (cf. text).

* under the heteroskedasticity specification $\sigma_i = c \cdot \exp(z_i \gamma)$; estimates of $\hat{c}$ and $\hat{\gamma}$ suppressed in table and available upon request. ** estimates of $\hat{\sigma}$ and $\hat{\rho}$ suppressed in table and available upon request.

Table 5a: Censored Regression Quantiles
— Baseline Specification, 1993 —
explained: share of riskless in financial wealth

θ	**0.1**		**0.2**		**0.3**		**0.4**	
	coeff.	t-value	coeff.	t-value	coeff.	t-value	coeff.	t-value
constant	2.6944	(3.185)	2.9966	(1.492)	6.6485	(0.081)	3.3911	(0.356)
age	–0.0643	(–1.458)	–0.0756	(–1.995)	–0.1372	(–2.213)	–0.1381	(–1.867)
age^2/100	0.1331	(1.424)	0.1436	(1.823)	0.2527	(1.968)	0.2536	(1.600)
age^3/1000	–0.0084	(–1.326)	–0.0081	(–1.535)	–0.0142	(–1.654)	–0.0141	(–0.657)
secondary educ.	–0.1687	(–1.798)	–0.0962	(–1.037)	–0.0877	(–0.760)	–0.0307	(–0.216)
pre–university	–0.1637	(–1.802)	–0.1045	(–1.314)	–0.1467	(–1.369)	–0.0880	(–0.776)
junior vocat.	–0.2344	(–2.359)	–0.1552	(–1.724)	–0.1575	(–1.404)	–0.0731	(–0.583)
senior vocat.	–0.2474	(–2.550)	–0.1679	(–2.015)	–0.1814	(–1.825)	–0.1425	(–1.175)
vocat. college	–0.1055	(–1.208)	–0.0408	(–0.562)	–0.0512	(–0.550)	0.0137	(0.117)
university	–0.1172	(–1.329)	–0.0564	(–0.766)	–0.0434	(–0.478)	0.0194	(0.167)
diploma	0.0447	(1.148)	0.0461	(0.985)	0.0491	(0.897)	0.0859	(1.755)
log fin. wealth 1	–0.1396	(–1.860)	–0.1180	(–0.545)	–0.4129	(–0.042)	–0.0035	(–0.003)
log fin. wealth 2	–0.1548	(–1.463)	–0.2813	(–3.076)	–0.1594	(–1.212)	–0.2669	(–2.434)
log fin. wealth 3	–0.1405	(–1.660)	–0.1130	(–1.280)	–0.1979	(–2.064)	–0.2175	(–2.034)
log fin. wealth 4	–0.1176	(–2.019)	–0.2033	(–3.441)	–0.2625	(–5.008)	–0.3098	(–4.475)
log fin. wealth 5	–0.0505	(–2.083)	–0.0904	(–3.446)	–0.1072	(–3.885)	–0.1416	(–4.581)
self–employed	–0.1038	(–1.999)	–0.0281	(–0.420)	–0.0098	(–0.171)	0.0187	(0.391)
no paid job	0.0749	(0.798)	0.0111	(0.118)	–0.0061	(–0.053)	–0.0577	(–0.394)
other job	–0.0150	(–0.164)	–0.0078	(–0.075)	0.0071	(0.051)	–0.0214	(–0.119)
retired / disabled	–0.0532	(–0.933)	–0.0370	(–0.692)	–0.0735	(–1.366)	–0.0642	(–0.907)
unemployed	0.2495	(1.522)	0.2254	(1.362)	0.1966	(1.165)	0.2437	(1.426)
married or cohab.	–0.0341	(–0.999)	–0.0011	(–0.022)	0.0202	(0.392)	0.0078	(0.127)
panel (1=HIP)	0.0076	(0.272)	–0.0350	(–1.126)	–0.0664	(–1.647)	–0.0760	(–1.657)
high income unc.	0.0621	(1.415)	0.1008	(1.986)	0.0870	(1.616)	0.0686	(1.111)
moder. inc. unc.	0.0469	(1.405)	0.0295	(0.672)	0.0584	(1.333)	0.0756	(1.426)
Nobs = 1915								
function	109.057		168.876		198.486		206.882	
Pseudo-R^2	0.1399		0.2443		0.2943		0.3101	
Wald, 2 df	2.412		4.062		2.690		2.078	

Note: Covariance matrix estimated by design matrix bootstrap estimator with 250 bootstrap iterations. $t-$values in parentheses; the Wald test statistics refer to joint significance tests of parameters for income uncertainty.

Table 5b: Censored Regression Quantiles
— Sensitivity Analysis, 1993 —
explained: share of riskless in financial wealth

θ	**0.1**		**0.2**		**0.3**		**0.4**	
	coeff.	t-value	coeff.	t-value	coeff.	t-value	coeff.	t-value
a) baseline specification, cf. table 5a — Nobs = 1915								
high income unc.	0.0621	(1.415)	0.1008	(1.986)	0.0870	(1.616)	0.0686	(1.111)
moder. inc. unc.	0.0469	(1.405)	0.0295	(0.672)	0.0584	(1.333)	0.0756	(1.426)
Pseudo-R^2	0.1399		0.2443		0.2943		0.3101	
Wald, 2 df	2.412		4.062		2.690		2.078	
b) 1 year horizon — Nobs = 1915								
high income unc.	0.0662	(0.808)	0.1429	(2.654)	0.1446	(2.382)	0.0338	(0.554)
moder. inc. unc.	0.0190	(0.678)	-0.0006	(-0.017)	0.0348	(0.971)	0.0158	(0.399)
Pseudo-R^2	0.1397		0.2470		0.2961		0.3095	
Wald, 2 df	0.958		8.879		5.673		0.314	
c) continuous measure of income variance — Nobs = 1765								
income variance	0.0010	(1.579)	0.0015	(1.909)	0.0007	(1.252)	0.0003	(0.475)
Pseudo-R^2	0.1481		0.2450		0.2918		0.3074	
d) precautionary motives, broad specification — Nobs = 1874								
prec. motive	-0.0010	(-0.052)	0.0334	(1.920)	0.0447	(2.230)	0.0407	(2.011)
Pseudo-R^2	0.1489		0.2565		0.3048		0.3184	
e) precautionary motives, narrow specification — Nobs = 1915								
prec. motive	-0.0004	(-0.012)	0.0434	(1.428)	0.0414	(1.346)	0.0654	(1.851)
Pseudo-R^2	0.1385		0.2423		0.2935		0.3115	
f) with interaction income expectation — Nobs = 1895								
high income unc.	0.0624	(1.244)	0.1196	(2.255)	0.0772	(1.416)	0.0793	(1.274)
moder. inc. unc.	0.0434	(1.252)	0.0416	(0.922)	0.0400	(0.860)	0.0731	(1.338)
income expect.	0.0004	(0.449)	0.0002	(0.130)	-0.0003	(-0.128)	-0.0006	(-0.175)
hi unc × expect.	-0.0005	(-0.236)	-0.0025	(-1.243)	-0.0024	(-0.912)	-0.0021	(-0.489)
mod unc × expect.	0.0006	(0.349)	0.0020	(1.068)	0.0019	(0.767)	0.0000	(0.001)
Pseudo-R^2	0.1411		0.2466		0.2964		0.3113	
Wald, 2 df	1.987		6.541		2.043		1.975	
Wald, 2 df (interact)	0.205		4.755		4.206		0.519	
g) interaction with log financial wealth — Nobs = 1915								
high income unc.	0.5242	(1.270)	0.1398	(0.404)	0.1898	(0.489)	−0.2088	(⋆)
moder. inc. unc.	0.1311	(0.402)	−0.1192	(−0.322)	0.2985	(0.761)	0.5575	(⋆)
hi unct × log wealth	−0.0417	(−1.155)	−0.0028	(−0.091)	−0.0099	(−0.000)	0.0274	(⋆)
mod unct × log wealth	−0.0076	(−0.271)	0.0144	(0.445)	−0.0213	(−0.000)	−0.0438	(⋆)
Pseudo-R^2	0.1424		0.2449		0.2943			
Wald, 2 df	1.685		0.836		0.597			
Wald, 2 df (interact)	1.453		0.501		0.106			

continued on next page

Table 5b: *continued*

θ	**0.1**		**0.2**		**0.3**		**0.4**	
	coeff.	t-value	coeff.	t-value	coeff.	t-value	coeff.	t-value
h) exclusion of log financial wealth — Nobs = 1915								
high income unc.	0.0331	(0.566)	0.1182	(1.719)	0.0960	(1.317)	0.1185	(1.445)
moder. inc. unc.	0.0073	(0.187)	0.0537	(0.934)	0.0179	(0.334)	–0.0025	(-0.041)
Pseudo-R^2	0.0334		0.0777		0.1101		0.1380	
Wald, 2 df	0.381		3.075		2.036		2.980	
i) prediction of median log financial wealth — Nobs = 1915								
high income unc.	0.0625	(1.283)	0.1030	(1.741)	0.1027	(1.271)	0.1159	(1.476)
moder. inc. unc.	0.0441	(1.186)	0.0699	(1.425)	0.0578	(1.180)	0.0234	(0.382)
Pseudo-R^2	0.0402		0.1001		0.1426		0.1786	
Wald, 2 df	1.969		3.088		1.890		2.922	
j) with pension entitlement — Nobs = 1915								
high income unc.	0.0410	(0.723)	0.1056	(2.011)	0.0830	(1.560)	0.0748	(1.263)
moder. inc. unc.	0.0320	(0.795)	0.0273	(0.677)	0.0591	(1.359)	0.0761	(1.396)
pension entitlem.	-0.0319	(-0.519)	-0.0502	(-0.843)	-0.0486	(-0.787)	-0.0531	(-0.733)
Pseudo-R^2	0.1384		0.2448		0.2939		0.3104	
Wald, 2 df	0.693		4.182		2.577		2.093	
k) with bequest motive — Nobs = 1878								
high income unc.	0.0669	(1.450)	0.1013	(2.098)	0.0838	(1.712)	0.0346	(0.525)
moder. inc. unc.	0.0511	(1.391)	0.0302	(0.717)	0.0633	(1.463)	0.0367	(0.616)
bequest motive	-0.0142	(-0.414)	0.0067	(0.149)	0.0424	(0.935)	0.0216	(0.570)
Pseudo-R^2	0.1359		0.2434		0.2945		0.3125	
Wald, 2 df	2.699		4.664		3.200		0.386	
l) with price expectation — Nobs = 1874								
high income unc.	0.0671	(1.275)	0.1043	(1.980)	0.0544	(0.829)	0.0382	(0.526)
moder. inc. unc.	0.0489	(1.224)	0.0408	(1.021)	0.0458	(0.908)	0.0492	(0.860)
price expectation	0.0023	(1.067)	0.0008	(0.379)	0.0002	(0.084)	-0.0012	(-0.490)
Pseudo-R^2	0.1467		0.2491		0.2983		0.3159	
Wald, 2 df	2.031		3.924		0.871		0.777	

Note: $t-$values in parentheses; the Wald test statistics refer to joint significance tests of parameters of interest — usually income uncertainty or precautionary motives, or interaction terms with other variables; covariance matrix: design boostrap estimator, 250 replications; $\star$ = calculation failed

Table 6a: Honoré Fixed Effects Estimator
explained: share of riskless in financial wealth

	1993/1994		**1993/1995**		**1994/1995**		**1993–1995**	
	coeff.	t–value	coeff.	t–value	coeff.	t–value	coeff.	t–value
log fin. wealth 1	–0.0958	(–1.675)	–0.1082	(–1.596)	–0.1782	(–4.295)	–0.1479	(–2.181)
log fin. wealth 2	–0.2463	(–3.067)	–0.1588	(–1.692)	–0.0994	(–1.507)	–0.1376	(–1.304)
log fin. wealth 3	–0.0505	(–0.668)	–0.0836	(–0.961)	–0.1949	(–2.774)	–0.1164	(–1.090)
log fin. wealth 4	–0.1980	(–3.275)	–0.2637	(–4.431)	–0.1962	(–3.300)	–0.2247	(–2.720)
log fin. wealth 5	–0.0905	(–1.786)	–0.1381	(–3.336)	–0.1634	(–4.103)	–0.1437	(–2.364)
high income unc.	0.0346	(0.955)	0.0505	(1.091)	0.0450	(1.135)	0.0370	(0.724)
moder. inc. unc.	0.0476	(1.633)	0.0853	(2.398)	0.0081	(0.262)	0.0412	(1.027)
Wald, 2 df	2.823		6.550		1.935		1.056	
Nobs (households)	1486		1041		1259		1840	
right cens. (%)	41.0		34.7		36.8			
function	104.278		100.195		107.926			
test for symmetry							79.78 (6df)	

Note: the Wald test statistics refer to joint significance tests of parameters for income uncertainty; the corresponding χ^2 value is: $\chi^2_2(95\%) = 5.991$

Table 6b: Honoré Fixed Effects Estimator
— Sensitivity Analysis, 1993–1995 —
explained: share of riskless in financial wealth

a) baseline spec., cf. table 6a — $N_H = 1840$		
high income unc.	0.0370	(0.724)
moder. inc. unc.	0.0412	(1.027)
Wald, 2 df	1.056	

b) 1 year horizon — $N_H = 1840$		
high income unc.	-0.0022	(-0.038)
moder. inc. unc.	0.0065	(0.195)
Wald, 2 df	0.064	

c) cont. measure of income variance, $N_H = 1665$		
income variance	-0.0008	(-1.261)

d)$\star$ precautionary motives, broad — $N_H = 1003$		
prec. motive	-0.0348	(-1.962)

e)$\star$ precautionary motives, narrow — $N_H = 1040$		
prec. motive	-0.0590	(-2.116)

f) with interact. inc. expect. — $N_H = 1665$		
high income unc.	0.0403	(0.747)
moder. inc. unc.	0.0391	(0.923)
income expect.	-0.0003	(-0.038)
hi unc $\times$ expect.	-0.0032	(-0.308)
mod unc $\times$ expect.	-0.0004	(-0.054)
Wald, 2 df	0.865	
Wald, 2 df (interact)	0.134	

g) with interact. log finan. wealth — $N_H = 1840$		
high income unc.	-0.2913	(-0.650)
moder. inc. unc.	-0.0961	(-0.254)
hi unct $\times$ log wealth	0.0308	(0.768)
mod unct $\times$ log wealth	0.0122	(0.369)
Wald, 2 df	0.583	
Wald, 2 df (interact)	0.703	

h) exclusion of log finan. wealth — $N_H = 1840$		
high income unc.	0.0351	(0.581)
moder. inc. unc.	0.0329	(0.688)
Wald, 2 df	0.485	

j) with pension entitlement — $N_H = 1841$		
high income unc.	0.0386	(0.759)
moder. inc. unc.	0.0418	(1.042)
pension entitlem.	-0.0305	(-0.364)
Wald, 2 df	1.086	

k)$\star$ with bequest motive — $N_H = 1040$		
high income unc.	0.0501	(1.100)
moder. inc. unc.	0.0855	(2.404)
bequest motive	0.0307	(0.838)
Wald, 2 df	6.544	

l)$\star$ with price expectation — $N_H = 1796$		
high income unc.	0.0317	(0.617)
moder. inc. unc.	0.0367	(0.931)
price expectation	0.0012	(0.531)
Wald, 2 df	0.870	

Note: N_H = number of observations per household; $t-$values in parentheses; the Wald test statistics refer to joint significance tests of parameters of interest — usually income uncertainty or precautionary motives, or interaction terms with other variables; estimates displayed are based on unbalanced panel data for 1993–1995, Minimum Distance estimates; estimates marked with $\star$ are two–wave Honoré estimates based on the years 1993 and 1995.

Appendix C: Optimization Algorithms for Censored Regression Quantiles

This Appendix gives a short account of optimization for censored regression quantiles, including censored least absolute deviations (CLAD).

To begin with, consider the (uncensored) model

$$y_i = x_i\beta_\theta + \epsilon_{\theta i} \tag{10}$$

for which the conditional quantile restriction

$$Q_\theta(\epsilon_{\theta i}|x_i) = 0 \tag{11}$$

holds. Koenker and Bassett (1978) introduce the estimator

$$b_\theta = \arg\min_{\beta_\theta} \sum_{i=1}^{n} \rho_\theta(y_i - x_i\beta_\theta); \qquad \rho_\theta(z) = (\theta - 1[z < 0])z. \tag{12}$$

Now, note that any least absolute deviation (LAD) problem (without censoring) has a representation as a linear program (LP; see for instance Bloomfield and Steiger (1983)). Then, (12) represents the primal problem of the LP. A global optimum exists if the matrix x has full column rank. It can be found by a pivoting algorithm which searches over a set of basic feasible solutions of a system of constraints (e.g. Barrodale and Roberts (1973)).

This method generalizes to other quantiles as well since the minimand for a quantile regression problem is inherently linear and convex in β: Define $u_i = \max\{\epsilon_i, 0\}$ and $v_i = -\min\{\epsilon_i, 0\}$ where $\epsilon_i = y_i - x_i\beta_\theta$. Let $b_j = \max\{\beta_j, 0\}$ and $c_j = -\min\{\beta_j, 0\}$. Thus, $\epsilon_i = u_i - v_i$ and $\beta_j = b_j - c_j$. The to (12) equivalent LP is

$$\begin{aligned}
\text{min} \quad & \sum_{i=1}^{n}(\theta u_i + (1-\theta)v_i) \\
\text{s.t.} \quad & y_i = \sum_{j=1}^{k} x_{ij}(b_j - c_j) + u_i - v_i \\
\text{and} \quad & b_j, c_j, u_i, v_i \geq 0 \quad \forall i = 1, \ldots, N, j = 1, \ldots, k
\end{aligned}$$

The equivalence of quantile regression and LP is destroyed as soon as censoring is taken into account. The model, as considered in Powell (1986a), specifies a linear relationship between the latent variable y^* and regressors x

$$y_i^* = x_i\beta_\theta + \epsilon_{\theta i} \tag{13}$$

where one only observes

$$y_i = \begin{cases} y_i^* & \text{if } \; y_i^* \leq yc_i \\ yc_i & \text{if } \; y_i^* > yc_i \end{cases}$$

Again, the conditional quantile restriction $Q_\theta(\epsilon_{\theta i}|x_i) = 0$ is imposed.

The censored quantile regression problem

$$\hat{B}_n(\theta) = \arg\min_{\beta} Q_n(\beta) = \sum_{i=1}^{n} \rho_\theta(y_i - \min\{yc_i, x_i\beta_\theta\}) \tag{14}$$

is not linear, due to the censoring (in our application at $yc_i = 1$ from above). Moreover, the function is non–convex and the solution $\hat{B}_n(\theta)$ need not be unique.

Fitzenberger (1994) develops an algorithm based on <u>B</u>arrodale–<u>R</u>oberts for <u>cens</u>ored quantile regressions (BRCENS), showing that (14) can be formulated "close" to an LP:

$$\begin{aligned} \min \quad & \sum_{i=1}^{n}(\theta u_i + (1-\theta)v_i) \\ \text{s.t.} \quad & y_i = \min\left\{\sum_{j=1}^{k} x_{ij}(b_j - c_j), yc_i\right\} + u_i - v_i \\ \text{and} \quad & b_j, c_j, u_i, v_i \geq 0 \quad \forall i = 1, \dots, N, j = 1, \dots, k \end{aligned}$$

where u_i, v_i, b_j and c_j are as above, but the argument ϵ_i of u_i and v_i is now $\epsilon_i = y_i - \min\{x_i\beta_\theta, yc_i\}$.

BRCENS will not necessarily find a global minimum due to the non–convexity of (14). However, with BRCENS we found the smallest function value in the application, compared to the results from various alternative algorithms.

Viable alternatives to BRCENS which are *designed* to solve (14) are Womersley (1986), and Koenker and Park's (1996) NLRQ. The latter proved inferior to BRCENS in the application. In addition Buchinsky (1994) proposed the following short–cut: it solves (12) for $\tilde{\beta}_\theta^{(j)}$ in iteration j using those observations for which in the *previous* iteration $x_i\tilde{\beta}_\theta^{(j-1)} < yc_i$. When in two consecutive iterations the set of observations are the same, the algorithm terminates using $\tilde{\beta}_\theta^{(j)}$ as minimzer of (14). This method is dubbed "iterative linear programming algorithm (ILPA)". We found difficulties with ILPA since it cycled in the sense that after a finite fixed number of iterations J, the sequence of observations used repeated itself. Fitzenberger (1994) shows that ILPA is based on theoretically untenable claims. This can lead not only to non–convergence but even conditional upon convergence does not guarantee a local minimum. Nevertheless it seems to work reasonably in practice as Fitzenberger's simulation exercises demonstrate. Our experience in the application confirms this only for the lower quantiles, where cycling was less serious.

Note, that many numerical minimizers which do not take the specific structure of (14) into account, will have trouble to find a global minimum of the objective function. This is because the objective function has many local minima due to the high dimension of the parameter space and the high degree of censoring. Gradient based methods, for instance, which are strictly speaking not applicable since (14) is not differentiable everywhere get easily trapped in local minima.

Powell (1984, 1986a) recommends M. J. D. Powell's directions set method and the downhill simplex method of Nelder and Mead (see Press et al. (1989) for reference and implementation) for low–dimensional problems. In our application we had difficulties in obtaining estimates with these algorithms.

From our experience emerges that LP based methods seem to be the only feasible, theoretically appealing, and computationally efficient ways to compute the estimates. Fitzenberger (1997) provides a neat presentation and thorough discussion of censored regression quantiles and available algorithms.

Appendix D: Covariance Matrix Estimation for Censored Regression Quantiles

Powell's (1984, 1986a) formulae for an estimate of the asymptotic covariance matrix differ according to whether the errors are i.i.d. or heteroskedastic, and involve the choice of a bandwidth parameter for a non–parametric estimate of the (conditional) error density. We experienced a rather high sensitivity of the covariance matrix estimate to the choice of this smoothing parameter. Hall and Horowitz (1990) propose a data–dependent, semi–automatic method for choosing both smoother (kernel) and bandwidth which are optimal in some sense. Their formulae involve in a first stage non–parametric estimates of first and second derivatives of a density function, based on two different bandwidth parameters. In a second stage this allows to estimate the bandwidth of interest. The Hall and Horowitz method is derived for the homoskedastic case, whereas our data are heteroskedastic. Therefore, Powell's estimator does not appear to be a wise choice to estimate the covariance matrix.

Alternatively, Buchinsky (1994) suggests to use the order statistics estimator of the covariance matrix, as developed by Huber (1981). The errors in the original model are assumed to be conditionally independent of the regressors; otherwise, the order statistics estimator will be inconsistent.

Finally, one can use bootstrap methods. The design matrix bootstrap estimator survives under heteroskedasticity and censoring as a consistent estimator. It is based on re–sampling observations with replacement. Hahn (1995) provides theoretical justification for using the bootstrap for Powell's (1986a) estimator. Buchinsky (1994, 1996), Chamberlain (1994), and Fitzenberger (1994, 1997) provide empirical and Monte Carlo evidence in favor of the bootstrap under heteroskedasticity.

The number of bootstrap samples drawn, B, is a 'smoothing parameter' to be chosen by the investigator. However, the estimate will become more accurate as $B \to \infty$, whereas the selected bandwidth for Powell's estimator may be either too large or too small. Andrews and Buchinsky (1996) devise a semi–automatic procedure for the choice of B for estimating standard errors. This choice depends in particular on the accuracy one wishes to obtain for the estimates (as measured by the relative deviation of the bootstrapped standard error estimate from the asymptotic value) and on the degree of excess kurtosis of the bootstrapped distribution of the parameter under consideration. With the Andrews and Buchinsky (1996) instructions for an accuracy parameter of 10% we found that for most of the parameters the number of required bootstrap estimates was well below 250 — including the parameters of particular interest. Some further experiments showed that increasing B to 500 or 1000 did not lead to different conclusions, at the expense of substantially longer calculation times.

We ignore the subtle point of asymptotic refinements of the bootstrap estimator to the empirical levels of asymptotically pivotal test statistics, as considered by Horowitz (1997). The reason is again that the bootstrapped distribution of, say, $t-$ or χ^2- statistics is based on some consistent estimator of the asymptotic covariance matrix in the first place. Horowitz develops a smoothed version of Powell's (1984) CLAD, "SCLAD", and documents that both slope coefficients and covariance matrix estimate are rather insensitive to the choice of bandwidth, hence allowing for bootstrapping $t-$ and χ^2- statistics.